University Of Canterbury

The potential for toxin and antitoxin gene pairs to display a post-segregational killing phenotype, with regards to the ecology of mobile elements.

Dorien Skye Coray

2014

A thesis submitted in partial fulfilment of the requirements
for the Degree of Doctor of Philosophy
in
Cellular and Molecular Biology

Abstract

Genes are able to replicate horizontally and vertically- a given gene may be more successful on horizontally mobile elements than others. This includes genes that exhibit a post-segregational killing (PSK) phenotype. PSK is generated by expression of a toxin and antitoxin from a mobile element, such that if a bacterium loses the element the toxin becomes active in the cell and the cell dies. All PSKs described to date involve a toxin and an antitoxin function, though within a given group of toxin and antitoxin gene pairs only some are likely to exhibit this phenotype. Here, I investigate what differentiates genes that induce PSK from biochemically similar genes that do not.

One group of genes of which some are known to induce PSK is toxin-antitoxin (TA) systems, composed of a stable toxin and an unstable antitoxin. I analyzed computational data on the distribution of type I TA systems (RNA antitoxin), which appear to be less mobile than type II TA systems (protein toxin). Data on validated TAs suggests a correlation between distribution, mobility and the PSK phenotype. Differences in phylogeny could be due to differences in tendency to exhibit PSK in different environments. This connection between distribution and PSK was explored by experimentally testing a computationally described operon, plasmid_Toxin-ptaRNA1, that exhibited structural and distributional similarities to a mobile type I TA system. Despite this, expression of the predicted toxin ORFs did not reduce growth (as measured by saturation density) in E. coli, and the operon did not induce PSK.

The conditions of PSK were further tested with the toxin (barnase) and antitoxin (barstar), which are not known to have the phenotype. A number of heterologous expression systems were developed with these genes in E. coli to test their ability to exhibit PSK in a manner akin to both type II TA systems, with a cytoplasmic toxin, and bacteriocins,

which have a secreted toxin. I used equations of logarithmic decay to model the necessary expression of the proteins in the cell and their rate of decay after plasmid loss to enable PSK. My results suggest there is likely to be an evolutionary trend toward TA systems with high expression levels of very unstable antitoxins. Secreted barnase was also tested experimentally for its ability to induce PSK similar to bacteriocins, which exhibit a PSK-like phenotype in monoculture by driving maintenance of the immunity encoding plasmid. Barnase did not induce PSK, possibly due to its inability to cause antibiosis in our test system.

Structural similarities and biochemical similarities are not sufficient to determine whether a given system will act as a PSK because numerous contextual factors have an effect on whether the genes are addictive. A given set of genes may have the phenotype in one species but not another, under one set of environmental conditions but not another, or on one replicon but not another. This is consistent with the competition hypothesis, which states that genes will be selected for on mobile elements due to their ability to increase horizontal reproductive success, depending on the environmental conditions.

Contents

Contents **v**

List of Figures **vii**

List of Tables **xi**

List of abbreviations **xiii**

Acknowledgements **xv**

1 Introduction **1**

1.1 Factors affecting gene mobility 2

1.1.1 Periodic selection 3

1.1.2 Infectious relatedness hypothesis 4

1.1.3 Selfish operon 4

1.1.4 Complexity hypothesis 4

1.1.5 Competition hypothesis 5

1.2 Genetic elements known to induce PSK 8

1.2.1 Toxin-antitoxin systems 9

1.2.2 Restriction modification systems 20

1.2.3 Bacteriocins 23

1.2.4 ABI Mechanisms 24

1.3 Exploring the parameter space of PSK 26

1.3.1 Part I: Distribution and functionality of computationally identified PSKs 28

1.3.2 Part II: Ability of a given toxin and antitoxin gene pair to induce PSK 30

2 Materials and Methods **35**
2.1 Bacterial Strains and Culture Conditions 35
2.2 Cloning 36
2.3 Bacterial Growth Assays 42
2.4 Protein and RNA Analysis 45
2.4.1 Transcription analysis 45
2.4.2 Translation analysis 46
2.5 Zones of Inhibition 46
2.6 PSK tests 48
2.7 Computational Analyses 49
2.7.1 Type I TA toxin family screen 49
2.7.2 Testing PT-PtaRNA1 as a type I TA system 51
2.7.3 Distribution of barstar and barnase on plasmids and viruses . 51
2.7.4 Programs and statistics 52

3 Distribution of type I TA systems across replicons, comparing between families and other TA types **53**
3.1 Results and Discussion 54
3.1.1 Structure-based searches for new TA families 55
3.1.2 Sequence-based searches for new TA families 57
3.1.3 Distribution of loci from known families 61
3.1.4 Phylogenetic distribution, PSK and presence on mobile elements 65
3.1.5 Factors that may affect expression of PSK from type I TA systems 68
3.1.6 Biases within the databases 69
3.2 Conclusion 70

4 Experimental analysis of candidate type I TA system PT-ptaRNA1 **73**
4.1 Results 74
4.1.1 Heterologous expression of candidate type I toxins under P_{BAD} promoter on a low copy number vector 75
4.1.2 Heterologous expression of candidate type I toxins on high copy number vectors 84
4.1.3 Testing PT-ptaRNA1 operons for PSK 91

4.2 Discussion 96
4.2.1 Plasmid_Toxin expression did not decrease *E. coli* growth . . 96
4.2.2 PT-ptaRNA1 operons did not exhibit PSK 98

5 Engineering expression systems for exhibition of PSK by a given toxin and antitoxin, using barstar and barnase as model proteins 101
5.1 Results 102
5.1.1 Expression systems with barnase under the P_{tac} promoter . . 102
5.1.2 Expression systems with barnase under the P_{BAD} promoter . 109
5.1.3 Relationship between protein expression and degradation necessary for PSK 115
5.2 Discussion 122
5.2.1 Various expression systems compatible with barstar and barnase, but may not be sufficient for a PSK effect 123
5.2.2 Level of protein expression dependent on respective stability of toxin and antitoxin, which may affect PSK evolution . . . 124

6 Exhibition of PSK by secreted toxin barnase 131
6.1 Results 132
6.1.1 Searches for barstar and barnase on mobile elements 132
6.1.2 Testing barstar and secreted barnase as PSK systems in *E. coli* 136
6.1.3 Antibiosis of neighboring bacteria from barnase secretion in *E. coli* 143
6.2 Discussion 148
6.2.1 Barstar and barnase are not widespread on plasmids and viruses 148
6.2.2 Secreted barnase is insufficient to drive the PSK phenotype . 149
6.2.3 Secreted barnase does not cause ZOIs within this test system 150

7 Conclusion 153
7.1 PSK is difficult to predict from gene structure and protein biochemistry 153
7.2 PSKs are context dependent 154
7.3 Some PSKs may be more efficient across different taxa than others . 156
7.4 The competition model supports this account of PSKs 158
7.5 Responsiveness of TA systems to stress could be selected for on mobile loci 160

7.6 Conclusion 162

Appendices **165**

A Formulas and Solutions **167**

A.1 Media 167

A.1.1 Luria Burtani Medium 167

A.1.2 RM medium 167

A.1.3 M9 medium 167

A.2 Media supplements 167

A.2.1 Amino acids 167

A.2.2 Inducers 168

A.2.3 Antibiotics 168

A.3 Solutions for protein extraction and electrophoresis 168

B Sequences **169**

B.1 Type I TA toxin amino acid sequences 174

B.1.1 Hok 174

B.1.2 FlmA 176

B.1.3 Fst 178

B.1.4 Ldr 179

B.1.5 ShoB 179

B.1.6 Ibs 179

B.1.7 TisB 180

B.1.8 Plasmid_toxin 181

B.1.9 TxpA 181

C Supplementary Data **183**

C.1 Chapter 3 183

C.1.1 Sequence logos for type I TA toxin families and aggregated HMMs 183

C.2 Chapter 5 190

C.2.1 Protein profiles of CSP01 and CSBs01 190

C.3 Chapter 6 192
C.3.1 Testing PSK by secreted barnase by culturing double-transformed E.coli 192

Bibliography **230**

List of Figures

1.1 General mechanism of post-segregational killing by type II TA systems on plasmids. 6
1.2 The intersection between subgroups of toxin and antitoxin gene pairs, and the addiction phenotype. 27
1.3 Hydrolysis of RNA by barnase. 32
1.4 Interactions between barnase and barstar. 32

3.1 Membrane orientation of representative type I toxins. 56
3.2 Alignment of putative type I toxin family loci. 59
3.3 Predicted base pairing between putative type I toxin mRNA and downstream antisense strand. 60
3.4 Replicon distribution of type I TA system loci. 63
3.5 Aggregated HMMs derived from type I TA system toxins. 64

4.1 Organization of synthesized plasmid_Toxin and PT-RNA1 operon sequences. 76
4.2 Induction of potential toxin plasmid_Toxin from *Actinobacter baumannii*, under the P_{BAD} promoter 78
4.3 Induction of potential toxin plasmid_Toxin from *B. pseudomallei*, under the P_{BAD} promoter 79
4.4 Induction of potential toxin plasmid_Toxin from *X. campestri*, under the P_{BAD} promoter 80
4.5 Induction of potential toxin plasmid_Toxin from *E. coli* OK137, under the P_{BAD} promoter 81
4.6 Induction of potential toxin plasmid_Toxin from *P. aeruginosa* plasmid pMATVIM-7, under the P_{BAD} promoter 82

4.7 Growth of *E. coli* bearing pBAD33, with and without arabinose treatment . . . 83
4.8 Transcription of synthesized candidate plasmid_Toxin genes in *E. coli* strain Rel606. . . . 85
4.9 Transcription of synthesized candidate plasmid_Toxin genes in *E. coli* strain Rel606. . . . 86
4.10 Induction of candidate toxin plasmid_Toxin from *P. aeruginosa* plasmid pMATVIM-7 under the Lac promoter in *E. coli* strain Rel606 on solid media. . . . 88
4.11 Induction of potential toxin plasmid_Toxin under the Lac promoter in *E. coli* strain DH5α, control and *P. aeruginosa* plasmid pMATVIM-7. 89
4.12 Induction of potential toxin plasmid_Toxin from *P. aeruginosa* plasmid pMATVIM-7 and *E. coli* OK137 in the reverse orientation under the Lac promoter in *E. coli* strain DH5α. . . . 90
4.13 Effect of potential PSK-inducing TA systems on bacterial growth under conditions of forced plasmid loss in *E. coli* strain MG1655. . . 93
4.14 Effect of potential PSK-inducing operons on bacterial growth under conditions of forced plasmid loss in *E. coli* strain Rel606. . . . 94
4.15 Transcription of candidate toxins from PT-RNA1 operons in different bacterial species, expressed in *E. coli* strain Rel606. . . . 95

5.1 Organization of barstar and barnase genes on plasmid pBB01. . . . 102
5.2 Optical density of *E. coli* culture expressing barnase and barstar under the P_{tac} promoter. . . . 103
5.3 CFUs of *E. coli* culture expressing barnase and barstar under the P_{tac} promoter. . . . 104
5.4 Resuscitation of *E. coli* growth after loss of barnase inducer, with barnase on a high copy number plasmid. . . . 105
5.5 Resuscitation of *E. coli* growth after loss of barnase inducer, with barnase on a low copy number plasmid. . . . 107
5.6 Intracellular protein profile of *E. coli* expressing barnase under the P $_{tac}$ promoter. . . . 108
5.7 Organization of barstar and barnase genes on plasmid pBB02. . . . 109

5.8 Optical density of *E. coli* culture expressing barnase under the P_{BAD} promoter. 110
5.9 Optical density of *E. coli* culture expressing barnase under the P_{BAD} promoter with various concentrations of arabinose. 111
5.10 Effect of barnase induction under the promoter P_{BAD} on *E. coli* growth. 112
5.11 Organization of barstar and barnase genes on plasmid pBB05. 113
5.12 Pre-induction of barstar reduces barnase toxicity. 114
5.13 Decay of antitoxin and toxin after plasmid loss, with subsequent PSK. 116
5.14 Changing the relative starting populations and rate of decay of antitoxin in respect to a fixed toxin. 120
5.15 Relationship between the relative stability and stoichiometry of toxin and antitoxin necessary for PSK generation. 125

6.1 Presence of extracellular barnase in the supernatant of CSBB01 . . . 137
6.2 Proportion of bacteria retaining barstar and barnase containing plasmids during serial culture without selection. 139
6.3 Growth of *E. coli* bearing temperature-sensitive plasmids, with and without barnase, under barnase-inducing conditions. 142
6.4 Effect of barstar and barnase on bacterial growth under conditions of forced plasmid loss in *E. coli*. 144
6.5 Zones of inhibition from *E. coli* secreting barnase from pBB02 145
6.6 Zones of inhibition from *E. coli* secreting barnase from pBB05 147

C.1 TxpA 183
C.2 Plasmid_Toxin 183
C.3 Hok 184
C.4 FlmA 184
C.5 Fst 185
C.6 Ldr 185
C.7 Ibs 186
C.8 ShoB 186
C.9 TisB 187
C.10 Aggregate of Hok and FlmA families 187
C.11 Aggregate of Fst and Ldr families 187

C.12 Aggregate of Ibs, TisB and ShoB 188
C.13 Aggregate of TxpA, plasmid_Toxin, Hok and FlmA families 188
C.14 Aggregate of Fst, Ldr, Ibs, TisB and ShoB 189
C.15 Aggregate of all families 189
C.16 Intracellular protein profile of cells expressing a *lacZ* fragment. . . . 190
C.17 Intracellular protein profile of barnase-containing cells, uninduced. . 191
C.18 Percentage of cells retaining +/- PSK plasmids during serial culture without selection. 193

List of Tables

1.1 Selected families of TA systems 12

2.1 Strain genotypes 35
2.2 Plasmids and cloning 37
2.3 Primer sequences 41

3.1 Scoring type I TA system toxins by membrane topology using TMHMM 55
3.2 Phylogenetic distribution of type I toxin families 62
3.3 Phylogenetic distribution of type II toxin families 62
3.4 Phylogenetic distribution of type III toxin families 63

4.1 Location and coordinates of plasmid_Toxin ORFs synthesized for experimental validation 75
4.2 Differences in growth between cultures induced and uninduced for PT expression in two *E. coli* strains MG1655 and Rel606, reported as p-values 77

6.1 Microbial RNases and their inhibitors known to interact with barstar and barnase 133
6.2 Presence of extracellular RNases and inhibitors on plasmids 134

B.1 Sequences used in Blastp searches of NCBI database. 173

List of abbreviations

AB	Antibiotic	Ng	Nanograms
ABI	Abortive infection system	Nm	Nanometer
Amp	Ampicillin	OD	Optical density
Ara	Arabinose	ORF	Open reading frame
ATP	Adenosine triphosphate	PAGE	Polyacrylamide gel electrophoresis
cDNA	Copy DNA	PCR	Polymerase chain reaction
CFU	Colony forming units	PSK	Post segregational killing
CM	Chloramphenicol	PT	Plasmid_Toxin
DNA	Deoxyribonucleic acid	RACE	Rapid amplification of cDNA ends
DNase	Deoxyribonuclease	RBS	Ribosomal binding site
dNTP	Deoxynucleotide triphosphate	RM	Restriction modification system
EMBL	European Molecular Biology Laboratory	RNA	Ribonucleic acid
HGT	Horizontal gene transfer	RNase	Ribonuclease
HMM	Hidden Markov Model	SDS	Sodium dodecyl sulphate
Hr	Hour	SE	Standard error
ICE	Integrating conjugative elements	Sec	Second
IPTG	Isopropyl β-D-1-thiogalactopyranoside	SI	Super integron
Kan	Kanamycin	SLH	S-layer homology
kDA	Kilo Dalton	TA	Toxin antitoxin system
LB	Luria burtrani	TCA	Trichloroacetic acid
MCS	Multiple cloning site	TMHMM	Transmembrane HMM
MGE	Mobile genetic element	tRNA	Transfer RNA
Min	Minute	TS	Temperature sensitive
mRNA	Messenger RNA	UTR	Untranslated region
NCBI	National Center for Biotechnology information	ZOI	Zones of inhibition

Acknowledgements

A few paragraphs in a thesis is a pitifully inadequate way to acknowledge the enormity of the support I have received in the last four and a half years within my professional and personal life. But nonetheless, inadequacy is perforce preferable to nothing at all. So here goes.

I would like to thank Jack, for nurturing my professional development with your scientific and humanitarian ideals, and your enthusiasm for lateral thinking. As my senior supervisor, your contribution to this work has been huge. I would like to thank Paul for offering up some of your data for me to play with and answering all of my questions on the methods behind the madness. I would like to thank Ant and Brigitta for your practical insights into my work, Leighton for humouring my rough attempts at modeling, and Ryan, for always knowing where random chemicals have been stored, and being my first go to for any problem. In every laboratory space in every country we've had to shuttle across in these last few years. I would like to thank everyone in the MolBio and Virus laboratory groups- I couldn't imagine what my PhD would have been without the sing-alongs, dancing, and outrageously inappropriate humour.

I would like to thank my Alaskan family, who have supported me from far away, and my New Zealand whanau who have made this place home. There has been plenty of drama these last few years, with earthquakes and illness and serious injury, and you've stood with me through it all. And finally I would like to thank Jasper, who foolishly suggested that I typeset my thesis in LaTex then, even more foolishly, did oh so much work setting it up for me. Thank you for putting up with my moodiness with a smile and for all the times you told me that, despite my own feelings on the matter, I was doing brilliantly and that everything was going to be okay.

Dooby dooby doo.

Chapter 1

Introduction

The prokaryotic gene pool includes a rich diversity of genetic elements. This diversity contributes to a complex ecology through the interactions of chromosomes, phage, plasmids, integrating conjugative elements, integrons, and transposons. At times, some of these elements replicate and transmit vertically during cell division. At other times these elements can mediate their replication and transmission horizontally, through infectious transfer. Horizontal transmission allows genes to move between cells (Zhang et al. *2003; Thomas and Nielsen 2005; Ochman* et al. *2000; Heinemann* et al. *1996; Heinemann* et al. *2000). The most prominent of the known pathways are transduction, transformation, and conjugation. Genes can thus potentially change both their genetic context, where and what type of replicon they are on, and cellular context, the type of and environment of their host. Such transfer can even defy taxonomic boundaries arising from nominal species categories (Heinemann and Sprague Jr 1989). I am interested in what affects the success of genes as they move into new genetic and cellular contexts. How do interactions with other genetic elements within the cell affect this? I investigate post-segregational killing (PSK) as a framework to address these questions.*

PSK is a lethal phenotype of some gene modules, where cells that lose the genes cease to grow. This is primarily seen during segregation of plasmids during cell division. Plasmids with related replication systems, known as incompatibility groups, compete to replicate within the cell. They can segregate to different daughter cells during cell division, resulting in asymmetric transfer (Heinemann 1998). When one of the segregating plasmids is +PSK, the cell not getting the +PSK plasmid dies, along with the competing plasmid. Plasmids are a kind of mobile genetic element (MGE), capable of reproducing independently of cellular reproduction. The ability

of the MGE to replicate at a rate or at times that differ from the cell creates the opportunity for the MGE to evolve traits that are deleterious to the host. The expression of PSK is one of those traits.

There are multiple types of genetic systems, primarily made up of encoded toxic and antitoxic functions, which exhibit PSK when on plasmids. These include toxin-antitoxin (TA) systems (Gerdes *et al.* 1986; Gerdes *et al.* 1986; Lehnherr *et al.* 1993; Weaver *et al.* 2009), restriction modification (RM) systems (Naito *et al.* 1995; Nakayama and Kobayashi 1998), bacteriocins (Inglis *et al.* 2013) and abortive infection (ABI) systems (Samson *et al.* 2013; Dy *et al.* 2014). Their ability to mediate competition for host space has contributed to their success on mobile elements, with many PSKs widespread across replicons and species.

Here, I primarily consider TA systems. TA systems were the first discovered PSK module (Gerdes *et al.* 1986). They consist of a protein toxin and a labile antitoxin that quickly degrades upon gene loss. When the antitoxin pool is sufficiently diminished, the remaining toxin in the cytoplasm can then act upon the cell and PSK occurs. I am interested in what differentiates TA systems that induce PSK from biochemically similar TA systems that do not. What does the distribution of known and predicted TA systems say about their proposed functionality as PSK modules? Can I predict whether a given toxin and antitoxin will act as a PSK in a given environment and cellular context? Here, I review different theories on what makes genes successful on an MGE. I then describe TA systems and other genetic modules known to cause PSK.

1.1 Factors affecting gene mobility

The mobility of a given gene is affected by certain physical factors, including how it is transferred (e.g., transformation, transduction, conjugation (González-Candelas and Francino 2012; Thomas and Nielsen 2005; Ochman *et al.* 2000)), the type of nucleic acids moved (linear, double stranded, circular, RNA, etc (Heinemann *et al.* 1996; Heinemann *et al.* 2000)), the presence of integrases, and sequence similarity between replicons, permitting homologous recombination (Yahara *et al.* 2012). Yet while any gene may move by the above mechanisms, various selective pressures affect which are maintained on new genetic elements and potentially moved into new hosts.

MGEs can be considered parasites (Cooper and Heinemann 2005; Rankin *et al.* 2010), capable of replicating infectiously. This is in contrast to genes or genetic elements that are specific to chromosomes. As with parasites, MGEs are subject to selection tradeoffs between virulence and mutualism, and horizontal and vertical reproduction (Harrison 2013). Plasmids can carry genes that are beneficial to the host (mutualism), and therefore increasing in number though increasing vertical transmission. But plasmid carriage also has a cost (virulence) (Rensing *et al.* 2002; Starikova *et al.* 2013; Mc Ginty and Rankin 2012). A plasmid that is more costly than beneficial will lower vertical reproduction of the host and thus itself, which must be made up for by a corresponding increase in horizontal transfer (Rankin *et al.* 2010). Studies of plasmid evolution have shown that during serial culturing experiments, the cost of plasmid carriage can lower over time (Dahlberg and Chao 2003; Lenski *et al.* 1998; Starikova *et al.* 2013).

The presence of some genes on mobile elements is primarily attributed to host-independent selection (Heinemann and Silby 2003; Cooper and Heinemann 2000; Jain *et al.* 1999; Jain *et al.* 2002; Kobayashi 1998). In the presence of antibiotics, for example, an antibiotic gene can increase the success of an MGE vertically and horizontally as it moves into new hosts (Heinemann and Silby 2003; Lenski *et al.* 1998). But that can be different than the selection for those genes to be retained on a mobile replicon initially. This has been described as a two-phase process of selection, first at the gene/replicon level and then at the host level (Heinemann and Silby 2003). A few factors that affect the mobilization of genes are discussed below, with particular emphasis on competition between mobile elements.

1.1.1 Periodic selection

It would be expected that a strictly host-beneficial plasmid-borne gene would eventually end up on the chromosome (Rankin *et al.* 2010; Toleman and Walsh 2011). Many genes are only beneficial to the host under certain conditions, as when challenged by antibiotics or heavy metals. If located on a chromosome, these genes might be lost during periods without the selective pressure (Rensing *et al.* 2002; Ellstrand *et al.* 1999). MGEs, then, may provide a reservoir of genes undergoing periodic selection. A given MGE could persist in a population at low levels, then sweep through it during strong selection (Rensing *et al.* 2002; Bergstrom *et al.* 2000).

1.1.2 Infectious relatedness hypothesis

In a screen of sequence data from 21 *E. coli* strains, Nogueira *et al.* (2009) found that secreted proteins have been frequently lost and gained from genomes and are frequently associated with mobile elements. Secreted 'goods' are frequently social in nature, involved in cooperation and conflict between bacteria. It has been suggested that mobility favors social behavior by increasing relatedness at the loci involved, regardless of host genome. HGT also offers a possible anti-cheating mechanism, re-infecting cells that lose the social trait (Rankin *et al.* 2010). Though the plasmid could be displaced by an incompatible plasmid lacking the social trait, this may still function as an anti-cheating mechanism at the cellular level (Rankin *et al.* 2010).

1.1.3 Selfish operon

The selfish operon model is based on observations of bacterial chromosomes- that genes tend to cluster into operons, a rare phenomenon in eukaryotes (Lawrence and Roth 1996; Demerec and Hartman 1959). Lawrence and Roth suggested genes that are grouped by function have an increased chance of successful transfer and functionality in a new host (Lawrence and Roth 1996). The organization of the genes, irrespective of the their function, could be considered 'selfish' because it promotes reproduction of the genes without necessarily benefiting the host in any way.

1.1.4 Complexity hypothesis

The complexity hypothesis attempts to explain the historic mobility of genes found on chromosomes based on the number of interactions required of the corresponding protein (Jain *et al.* 1999; Jain *et al.* 2002). In essence, those genes required for housekeeping functions (biosynthesis genes, membrane proteins, intermediate metabolism) are more likely to have been acquired through HGT than informational genes (transcription, translation, macromolecular synthesis) (Rivera *et al.* 1998). The average subunit of a protein engaged in translation, an 'information' function, interacts with four to five other ribosomal gene products, while housekeeping proteins may interact with only one or only a substrate molecule (Jain *et al.* 1999). The more distant a particular gene from the phylogeny of a chromosome the less likely it will have retained the structural features that optimize its interactions with other proteins in the same cell.

Jain *et al.* (1999) suggests that because housekeeping genes interact with fewer other gene products, they require less to function in a new environment and thus are more likely to be successful than informational genes when moved into new hosts. This hypothesis says nothing about the frequency of gene transfer per se, but about the frequency at which genes are retained after transfer. That frequency is determined by both the innate function of the gene and the gene product's ability to integrate into multi-protein complexes. Both the complexity hypothesis and the selfish operon are based on the ability of the genes to function in a new host, but explain gene mobility regardless of what is most beneficial to the host.

1.1.5 Competition hypothesis

The competition model suggests that traits on MGEs can be selected because of their ability to increase the success and competitiveness of interacting elements, independent of any benefit to the host. Even mechanisms that do not directly increase reproduction (Cooper and Heinemann 2000), but only eliminate a competitor, can be selected. Replicons can compete for space and reproductive success within cells (Heinemann and Silby 2003). Many mobile elements, including plasmids, phage, and even individual gene modules, exhibit some form of incompatibility, where similar elements cannot coexist in the same cell (Nakayama and Kobayashi 1998; Velappan *et al.* 2007; Taylor and Grant 1976). Incompatible plasmids, for example, tend to segregate to different daughter cells during cell division. This is through shared replication and copy control mechanisms (Wang *et al.* 2009; Velappan *et al.* 2007; Ebersbach and Gerdes 2005), and results in the plasmid inhabiting one less cell than it would in the absence of a competitor. The elimination of a horizontally mobile competitor thus increases the reproductive fitness of an MGE, independent of its affect on host (vertical) reproduction. Different phenotypes of MGEs have been implicated in competition between replicons, including post-segregational killing (PSK). PSK causes cell death upon gene loss. It was first identified as a phenotype of toxin-antitoxin (TA) systems in the 1980's (Jaffe *et al.* 1985; Gerdes *et al.* 1986; Gerdes *et al.* 1986). Generally, TA systems are two gene modules, one of which encodes a stable protein toxin and one a labile antitoxin function. Should a cell lose its TA genes due to events such as plasmid missegregation or displacement by an incompatible plasmid (Cooper and Heinemann 2000; Cooper and Heinemann 2005),

the antitoxin quickly degrades, leaving the more stable toxin in the cytoplasm longer than the antitoxin (Figure 1.1). This results in the PSK phenotype, where the cell dies upon gene loss. PSK is often considered a type of addiction system, in that the cell becomes 'addicted' to the presence of the antitoxin in the module.

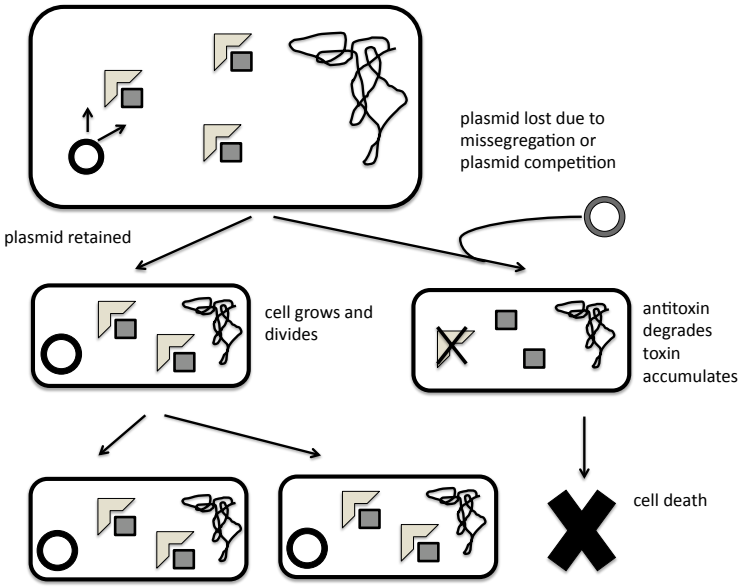


Figure 1.1: General mechanism of post-segregational killing by type II TA systems on plasmids. Within the cell, the TA genes are translated into a toxin and an antitoxin. If the plasmid is retained, there is a continuous source of antitoxin and the cell grows and divides. If the plasmid is lost due to missegregation or competition with an incompatible plasmid, the source of antitoxin is lost. The remaining antitoxin in the cytoplasm degrades, leaving the more stable toxin to accumulate. This results in cessation of cellular growth.

Testing the competition model

PSK was originally believed to have evolved as a mechanism of plasmid stability, ensuring transmission of the plasmid vertically (Gerdes *et al.* 1986), as opposed to having a role in the plasmid's horizontal lifestyle. This was tested in a series of experiments using the type II TA system ParDE and type I system Hok-Sok (Cooper and Heinemann 2000; Cooper and Heinemann 2005). If PSK were primarily

a mechanism of vertical stability, analogous to partitioning systems that ensure plasmid copies segregate to daughter cells during cell division, one would expect PSK to increase the number of copies of the plasmid in monoculture.

To test this Cooper and Heinemann (2000) infected *Escherichia coli* with either a plasmid that could cause PSK (+PSK) or an isogenic relative that could not (–PSK). Cultures were grown for 200 generations in separate monocultures. They found that the percentage of cells within the culture that retained the +PSK plasmids was higher than the percentage that retained the –PSK plasmids. This is consistent with mechanisms of PSK. Yet the total number of cells retaining the plasmid was the same in both cultures, showing that PSK modules do not increase vertical reproduction of a plasmid. When + and –PSK plasmid-bearing cells were competed under conditions of vertical reproduction, the percentage of +PSK cells went down in the population (Cooper and Heinemann 2005). This was probably due to the additional cost of carrying the PSK genes, and further demonstrates that PSK is not beneficial for vertical reproduction, a prediction of the plasmid stability model of PSK evolution (Cooper and Heinemann 2000). This indicated that addiction alone cannot compensate for the cost of PSK carriage, and is therefore unlikely to be the cause of TA systems being successful on MGEs.

Competition, on the other hand, allows for something to be beneficial without increasing absolute reproduction by eliminating competitors (Heinemann and Silby 2003). Indeed, the greatest benefits of PSK are seen under conditions involving HGT and competition between incompatible plasmids. When the incompatible plasmids segregate to daughter cells, the cells loosing the +PSK plasmid are susceptible to the released toxin, killing the cells as well as the competing –PSK plasmid (Naito *et al.* 1995; Cooper and Heinemann 2000). Experiments have shown that +PSK plasmids can resist displacement by incompatible –PSK plasmids and can successfully invade hosts with the –PSK plasmid (Cooper and Heinemann 2000; Naito *et al.* 1995; Nakayama and Kobayashi 1998). The same +PSK plasmids that were outcompeted during vertical reproduction rapidly took over the culture, by three orders of magnitude after 15 generations, when tested in cellular backgrounds that allowed for horizontal gene transfer.

This supports the competition model, where PSK is selected for during competition within host cells. This invokes within-host selection, seen in the evolution of parasite virulence (Nowak and May 1994) and bacteriophage infections (Turner

and Chao 1998). In the latter, success was shown to be dependent not on absolute rates of reproduction but on phage-phage intrahost competition. These results also fit with the selfish gene hypothesis, that PSKs kill plasmid-free competitors to aid remaining hosts and thus itself (Naito *et al.* 1995). But the selfish gene hypothesis still relies on vertical transmission, whereas the competition model selection is driven by plasmid-plasmid interactions, and explains the preferable success of PSK under horizontal conditions (Cooper and Heinemann 2000; Heinemann 1999; Peters and Benson 1995).

It is also possible that the death of plasmid-free cells would free up nutrients for plasmid-containing cells. Cooper and Heinemann (2000) found that in unstructured environments created in liquid culture, extra nutrients alone were not enough to make up for the cost of PSK. The role of freed nutrients or space may be more important in a low nutrient or structured environment (Cooper and Heinemann 2005; Rankin *et al.* 2010). The benefit of extra nutrients will be greater when resources are scarce, and in structured environments the available space/nutrients released by killing competitors are more likely to be available to related +PSK cells (Rankin *et al.* 2010).

The competition model can also explain the presence of some bacterial virulence and antibiotic resistance genes on plasmids (Eberhard 1990; Boyd *et al.* 2001; Amábile-Cuevas and Heinemann 2004). These are functionally PSK mechanisms in that they can render a dispensable plasmid essential in some environments (Cooper and Heinemann 2005). For example, cells can be said to be 'addicted' to an antibiotic resistance gene in the presence of the antibiotic, as cells that lose the genes die. Widespread enrichment for resistance mechanisms in clinical environments attests to the success of addiction genes. I am interested in what types of genes are successful on mobile elements. The ubiquitous nature of PSKs on mobile elements suggests they are one such type.

1.2 Genetic elements known to induce PSK

There are a number of systems that exhibit PSK. Though this work focuses on TA systems, I am interested in what differentiates TA systems that induce PSK from biochemically similar TA systems that do not. In attempting to narrow down those factors, it is helpful to examine other systems that exhibit PSK. This includes restriction modification systems, bacteriocins, and abortive infection systems.

1.2.1 Toxin-antitoxin systems

TA systems were the first identified PSK modules. They are currently grouped into five classes, all with a protein toxin but differing in the mechanism of antitoxin action (Table 1.1). The antitoxin can be an RNA or a protein, interacting directly or indirectly with the toxin. The recent proliferation of TA system 'types' may indicate a need for a pause and consideration of the classification and the meaning of TA system. The original designation of TA systems, namely types I and II, was tied to the ability of the genes to create the PSK phenotype on plasmids (Gerdes *et al.* 1986; Jaffe *et al.* 1985). Later, homologous or structurally similar pairs found on chromosomes with similar operon organization and regulation were considered TA systems.

Currently, the trend has been to generalize the label further, with many newly discovered gene pairs including a toxin and an antitoxin described thus. Such a broad definition is in danger of losing biologically relevant meaning. It remains to be seen whether TA systems form a group with a conserved or related biological function or if TA systems are a catalogue of all possible biochemical activities that when unbalanced cause toxicity. In large measure this thesis is concerned with that distinction. I propose a definition of canonical TA systems that includes organization into an operon, antitoxin-mediated regulation of toxin transcription (type II) or translation (type I), and high lability of the antitoxin relative to the toxin (Gerdes *et al.* 1986; Weaver *et al.* 2009; Sat *et al.* 2001; Christensen and Gerdes 2003; Johnson *et al.* 1996; Ogura and Hiraga 1983). This remains a structural definition rather than a functional definition. I review all five currently termed subtypes of TA systems in this chapter, but for the rest of the thesis will mostly consider types I, II, and III that adhere to the structural definition above.

Type I systems use an antisense RNA antitoxin. The toxins are mostly small, hydrophobic proteins that destabilize cellular membranes at high concentrations, though the exact mechanism of toxicity is not always known (Table 1.1) (Fozo *et al.* 2008; Gerdes *et al.* 1986; Ono *et al.* 1986; Weaver *et al.* 2003). SymE, a nuclease, is an exception (Kawano *et al.* 2005; Kawano *et al.* 2007). The RNA antitoxins are commonly encoded on the opposite strand as the toxin. This enables it to bind the toxin mRNA, blocking binding sites necessary for translation (Darfeuille *et al.* 2007; Kawano *et al.* 2007; Fozo *et al.* 2008; Fozo *et al.* 2008) and/or inducing RNase

degradation (Gerdes *et al.* 1992; Gerdes and Wagner 2007).

A well-studied example of a type I system is Hok-Sok, first identified on the R1 plasmid of *E. coli* (Gerdes *et al.* 1986; Ono *et al.* 1986). In the cell, the stable Hok mRNA is slowly processed at the 3' end into a translatable isoform (Franch *et al.* 1997). The antitoxin Sok, a 64-nucleotide antisense RNA, forms a duplex with this isoform of Hok mRNA that is rapidly degraded by RNase III (Kawano 2012; Gerdes and Wagner 2007). Sok is transcribed at a high rate but only has a half-life of 30 seconds, 1/40th of the Hok mRNA half-life (Franch *et al.* 1997). Sok rapidly disappears upon loss of the *sok* gene. In the absence of Sok, the remaining unprocessed Hok mRNA in the cytoplasm is able to mature and be translated into a toxic protein (Hayes 2003; Franch *et al.* 1997; Gerdes *et al.* 1990; Gerdes *et al.* 1992).

Type II TA systems utilize a protein antitoxin that directly binds the toxin. Many type II toxins characterized thus far target DNA replication (e.g., KidA, CcdB, ParE) (Hayes 2003), protein synthesis (Correia *et al.* 2006) or cell division (Masuda *et al.* 2012) (Table 1.1). The two genes are typically arranged in operons with the antitoxin gene preceding the toxin gene (Hayes 2003). Operon transcription is autoregulated by the antitoxin itself or in complex with the toxin (Feyter *et al.* 1989; Tam and Kline 1989; Magnuson and Yarmolinsky 1998; Cataudella *et al.* 2012), though other gene products can be involved (Hayes 2003).

Type II antitoxins tend to have largely disordered structures, that are highly susceptible to proteolysis compared to the toxins (Makarova *et al.* 2009; Buts *et al.* 2005). Loss of the antitoxin gene results in differentially rapid loss of antitoxin. In some instances toxicity is reversible upon addition of antitoxin (Gerdes and Wagner 2007; Buts *et al.* 2005; Christensen-Dalsgaard and Gerdes 2006; Pedersen *et al.* 2003). For example the TA system *ccd*, has reversible toxicity. The toxin CcdB stalls DNA gyrase by entrapping a cleavage complex between gyrase and DNA, forming a roadblock for DNA replication and transcription. The antidote CcdA competes with gyrase for CcdB binding, and addition of it can remove CcdB and allow gyrase to perform as usual (Maki *et al.* 1996; Melderen *et al.* 1994).

As the number of known type II loci has increased with recent computational searches, it has become clear that a given family of toxins can pair with a number of different antitoxin families (Leplae *et al.* 2011; Chopra *et al.* 2013; Makarova *et al.* 2009), making family designations based on the gene pair difficult. The known

type II systems were recently classified into 12 toxin and 20 antitoxin superfamilies (Leplae *et al.* 2011), based on shared protein domains. For example, the toxins of the MazF/PemK/CcdB superfamily share a ribonuclease SH3 fold and have several different types of cognate antitoxins, and RHH domain containing antitoxins are found together with ParE/RelE and CcdB/MazF type toxins (Leplae *et al.* 2011). Leplae *et al.* (2011) propose, from this, that toxin and antitoxin families have different ancestral origins and have been assembled or swapped multiple times during evolution.

Type III TA systems have only recently described (Fineran *et al.* 2009; Blower *et al.* 2012; Samson *et al.* 2013) (Table 1.1). The RNA antitoxin binds the protein directly, and appears to assemble into a triangular complex of three peptides of toxin interspersed by three RNA pseudoknots (Blower *et al.* 2011; Short *et al.* 2013). The originally described system consisted of the toxin ToxN and the antitoxin ToxI (Fineran *et al.* 2009). ToxN is structurally similar to the MazF/CcdB superfamily of type II toxins, and causes cell death upon overexpression, probably through ribonuclease activity (Blower *et al.* 2011; Short *et al.* 2013). Using ToxN as a seed for structure-based similarity searches, 125 putative type III operons were identified on plasmids and chromosomes of bacteria with a wide range of lifestyles (Blower *et al.* 2012).

Type IV and V TA systems have also been described only recently (Table 1.1). The type IV system was originally classified as a type II system, made up of a protein toxin and protein antitoxin. Yet it differs in that the antitoxin does not directly interact with the toxin (Masuda *et al.* 2012). The toxin YeeV inhibits polymerization of bacterial cytoskeletal proteins (Masuda *et al.* 2012; Tan *et al.* 2011). The antitoxin, the YeeU protein, binds the target cytoskeletal proteins and protects them from YeeV (Masuda *et al.* 2012). Like other TA systems, the antitoxin YeeU is preferentially degraded. The gene *yeeU* has been found both on chromosomes and in an integrated mobile element (Masuda *et al.* 2012; Brown and Shaw 2003). The *abiE* gene system has been shown to act like a type IV TA and can stabilize plasmids, suggesting the potential for these systems to exhibit PSK (Dy *et al.* 2014). The antitoxin AbiEi represses transcription of the *lacZ* gene under the AbiE promoter, the first evidence that type IV systems are negatively autoregulated similarly to previous TA systems (Dy *et al.* 2014).

The type V TA system involves a protein antitoxin, GhoS, that specifically cleaves the toxin mRNA, GhoT (Wang *et al.* 2012; Wang *et al.* 2013). GhoT is a membrane

lytic peptide (Wang *et al.* 2012). Unlike other protein TA systems, GhoS is not less stable than GhoT and does not regulate expression of the TA operon (Wang *et al.* 2012). The *ghoST* genes have only been found on chromosomes - it is unknown if the system acts as a PSK.

Table 1.1: Selected families of TA systems

Family	Toxin Activity	References
Type I (RNA antitoxin interacting with toxin mRNA)		
Hok-Sok	Inserts into inner membrane of the cell	(Gerdes *et al.* 1986; Gerdes *et al.* 1986; Ono *et al.* 1986)
TisB-IstR	Inserts into inner membrane of the cell	(Vogel *et al.* 2004; Unoson and Wagner 2008)
Fst-RNAII	Predicted transmembrane protein, believed to insert into inner membrane of the cell	(Weaver *et al.* 2009; Weaver 2012; Patel and Weaver 2006; Kwong *et al.* 2010)
SymE-SymR	Cleave mRNA, interfering with translation	(Kawano *et al.* 2005; Kawano *et al.* 2007)
Type II (Protein antitoxin interacting with protein toxin)		
MazEF	Cleaves mRNA, interfering with translation	(Sat *et al.* 2001; Zhang *et al.* 2003; Zhang *et al.* 2003)
RelBE	Cleaves mRNA, interfering with translation	(Christensen and Gerdes 2003; Neubauer *et al.* 2009; Pedersen *et al.* 2003)
PhdDoc	Inhibits ribosome, interfering with translation	(Liu *et al.* 2008)
ParDE	Inhibits gyrase, blocking chromosome replication	(Jiang *et al.* 2002; Fiebig *et al.* 2010)
HipBA	Phosphorylates elongation factor EF-Tu, interfering with translation	(Schumacher *et al.* 2009)
CcdAB	Inhibits gyrase, blocking chromosome replication	(Afif *et al.* 2001)
Type III (RNA antitoxin interacting with protein toxin)		
ToxIN	Probable ribonuclease	(Fineran *et al.* 2009; Blower *et al.* 2011; Short *et al.* 2013)
Type IV (Protein antitoxin interacting with target of protein toxin)		

Table 1.1: Selected families of TA systems continued...

Family	Toxin Activity	References
YeeUV	Inhibits polymerization of bacterial cytoskeletal proteins	(Masuda *et al.* 2012; Tan *et al.* 2011)

Distribution

TA systems have been found on plasmids (Gerdes *et al.* 1986; Poulsen *et al.* 1989; Weaver *et al.* 2009; Kwong *et al.* 2010), integrating conjugative elements (ICEs) (Wozniak and Waldor 2009), super integrons (SI)(Christensen-Dalsgaard and Gerdes 2006), phage (Guerout *et al.* 2013) and chromosomes (Pandey and Gerdes 2005). Many TA systems on chromosomally integrated mobile elements are still expressed. TA systems are also found on the chromosome in most bacterial and archaeal species (Pandey and Gerdes 2005), often many: the *E. coli* K-12 genome has at least 36 (Christensen and Gerdes 2003) and *Mycobacterium tuberculosis* 88 (Ramage *et al.* 2009). TAs appear to have moved frequently between chromosomes and plasmids, with chromosomal TAs often on genomic islands (Ramage *et al.* 2009). An exhaustive search for homologues of the MazF toxin found that plasmid-borne loci were interspersed between chromosomal modules, suggesting frequent interchange between plasmids and chromosomes (Chopra *et al.* 2013).

Though TAs are found on chromosomes of bacteria with diverse lifestyles and genome size (Pandey and Gerdes 2005; Mruk and Kobayashi 2014), the highest numbers of type II TAs have been found on slow growing organisms living under nutrient-limiting conditions (Mruk and Kobayashi 2014). According to recent bioinformatic work (referenced but unpublished), the prevalence of type II TAs is independent of chromosome size (Goeders and Van Melderen 2014). Many chromosomal TAs appear to no longer function: five of six copies of *hok-sok* on the *E. coli* K12 chromosome have insertions or point mutations, and none demonstrated PSK activity (Pedersen and Gerdes 1999; Gerdes and Wagner 2007).

There also appears to be differences in the distribution between types of TA systems. Thus far, type II systems appear to be more broadly dispersed than type I systems. A type I family is often restricted to a single phyla (Fozo *et al.* 2010), while many type II families occur in all bacterial phyla as well as some archaea

(Leplae *et al.* 2011; Makarova *et al.* 2009). In addition, type II families tend to have more evidence of HGT. They are commonly found on mobile elements, and are often distributed in a lineage independent manner, with loci appearing between clades without necessarily appearing in most species within the clade (Makarova *et al.* 2009; Leplae *et al.* 2011). Some type I families have only been described on chromosomes thus far (Fozo *et al.* 2010). The reason for these differences is unknown (Mruk and Kobayashi 2014).

Proposed functions of chromosomal TA systems

Two different speculations about the biological roles of chromosomal TAs predominate in the literature. One speculation is that TA systems are important components in cellular function. The other is that they are genomic 'parasites' that persist due to their addictive properties. The literature on this topic is expansive (and debated), but some major theories are evaluated here. The majority of work has focused on type II systems, as is reflected in the discussion below, though analysis of type I systems has been increasing.

The first two theories involve properties similar to plasmid TAs: reducing loss of genomic regions, particularly integrated mobile elements, and counteracting the PSK affect of TAs on mobile elements. The other three are all based on integration of the TAs into cellular networks, primarily those involved in stress response. The general stress response theories come from observations of chromosomal TA activation (toxin release) upon stress, particularly amino acid starvation. Two more specific forms of stress have also been linked to TA systems, and are discussed here: resistance of individual cells to antibiotics (persistence) and regulation of *Myxcoccus xanthus* sporulation.

Stabilization of genomic regions

Because of their addictive characteristics, chromosomal TA loci could stabilize genomic regions in a way analogous to their role on plasmids, inhibiting large-scale deletion of adjacent regions (Van Melderen and De Bast 2009; Szekeres *et al.* 2007; Yamaichi *et al.* 2007; Rowe-Magnus *et al.* 2003). Many chromosomal loci are in fact on integrated elements, associated with integrated conjugative elements, transposons and prophages. ICEs exist in equilibrium between the (predominant) chromosomal

and circular, extrachromosomal form (Wozniak and Waldor 2009). The type II TA locus *mosAT* on an ICE in *Vibrio cholera* appears to increase the stability of the ICE in circular form: the genes are de-repressed by the same proteins that promote excision (Wozniak and Waldor 2009). Super integrons are gene capture systems composed of a site-specific recombinase followed by numerous gene cassettes which can be excised, integrated, or shuffled (Rowe-Magnus and Mazel 2001). While SI cassettes are mobile, the platform of SIs appears to be sedentary within the chromosome, and many are particularly refractory to deletion (Rowe-Magnus *et al.* 2003). Two type II higBA loci from *V. cholera* SI were found to increase the stability of an unstable plasmid (par deficient R1 derivative) in monoculture by a factor of over 25-fold over 80 generations (Christensen-Dalsgaard and Gerdes 2006). Addition of two type II TA loci, *relBE* (mRNA cleavage) and *parDE* (gyrase inhibitor), isolated from a *Vibrio vulnificus* SI reduced gene loss in unstable SIs and a dispensable genomic region of *E. coli* K12 (Szekeres *et al.* 2007).

TA systems could also maintain regions of the chromosomes not associated with mobile elements (Szekeres *et al.* 2007). The maintenance of genomic regions regardless of selective pressure could be beneficial to bacteria by providing a reservoir of potentially adaptive material (Szekeres *et al.* 2007). It is difficult to say if this is a major function for chromosomal loci, though, as many are no longer able to stabilize plasmids and/or on chromosomal regions (Szekeres *et al.* 2007; Christensen *et al.* 2004; Wilbaux *et al.* 2007; Fiebig *et al.* 2010).

Anti-addiction

Another hypothesis is that chromosomal antitoxin could protect the cells from PSK when infected by plasmids with cognate toxins. By providing a continual source of antitoxin, any residual toxin remaining after loss of the plasmid would be neutralized. This hypothesis has experimental support. When the *parDE* operon was inserted into the *E. coli* chromosome, the benefit of PSK during plasmid competition was neutralized (Cooper and Heinemann 2000). Mine *et al.* (2009) found that cells containing a chromosomal type II *ccdBA* operon were immune to moderate levels of CcdB toxin expression from plasmids. The *Erwinia chrysanthemi* chromosomal TA ccd_{Ech} was found to both protect cells from PSK induced by the F plasmid *ccd* homologue (ccdF) and gave a competitive advantage over cells lacking the operon,

under conditions where PSK can be measured (De Bast *et al.* 2008). That said, many systems do not appear to cross-talk (Wilbaux *et al.* 2007; Fiebig *et al.* 2010; Grady and Hayes 2003). This may be due to divergence in protein structure, or low levels of expression from some chromosomal systems. Christensen *et al.* (2006) found that two type II *higBA* loci from *Vibrio* SI did not crosstalk, with toxins only 26% similar and antitoxins only 17% similar (Christensen-Dalsgaard and Gerdes 2006). The chromosomal type II TA *chpB* could neutralize plasmid TA toxin ParD, but only once the antitoxin was up-regulated (Santos Sierra *et al.* 1998).

This effect only requires that the antitoxin remains functional- the toxin in the operon is unnecessary. In fact, antitoxin 'orphans' are more common on chromosomes than toxin orphans (Mine *et al.* 2009; Mruk and Kobayashi 2014). This would be expected for inactivation of the TA system over time in general, as loss of the antitoxin in the presence of a functional toxin would cause cell death (Mine *et al.* 2009). But the antitoxin may also be selected for in the presence of toxins on mobile elements (Cooper and Heinemann 2000; Rankin *et al.* 2012). Rankin *et al.* (2012) suggested that a rock-paper-scissor mechanism might explain the presence of antitoxins on host chromosomes. The presence of an antitoxin on the host chromosome would allow the less costly –PSK plasmid to proliferate. Once the –PSK plasmid is well-established, the benefit of a chromosomal antitoxin would be reduced, allowing wild-type cells to replace antitoxin-containing cells. This sets the stage for the +PSK plasmid to invade again (Rankin *et al.* 2012).

General stress response

The general stress-response hypothesis is that the balance of toxin and antitoxin is perturbed during cellular stress response, releasing toxin and leading to bacteriostasis or, some argue, programmed cell death. Unlike plasmid-induced PSK, the chromosomal antitoxin gene is still present and could potentially rescue cells if expression resumed. It has been hypothesized that this could provide an adaptive advantage to the cell, transiently reducing growth in low nutrient/stressful environments (bacteriostasis), or providing food for living bacteria (programmed cell death) (Buts *et al.* 2005).

This theory primarily comes out of work done with chromosomal type II systems in *E. coli*, particularly MazEF (Engelberg-Kulka *et al.* 2006; Amitai *et al.* 2004; Sat

et al. 2001) and RelBE (Gerdes *et al.* 2005; Christensen *et al.* 2001; Christensen and Gerdes 2003). Many of the type II toxins affect macromolecule synthesis (Table 1.1), and overexpression of toxins in unstressed cells correlated with drops in translation rates (Pedersen *et al.* 2003; Christensen-Dalsgaard and Gerdes 2006; Christensen and Gerdes 2003) and DNA replication (Pedersen *et al.* 2003). This effect is often reversible upon over-expression of the cognate AT, suggesting that the cells are in a bacteriostatic state (Christensen-Dalsgaard and Gerdes 2006; Pedersen *et al.* 2003). Not all growth inhibition effects are reversible though: while many assays were done over a three to four hour range (Christensen-Dalsgaard and Gerdes 2006; Pedersen *et al.* 2003), Amitai *et al.* (2004) showed that the effects of expression of type II toxin MazF were irreversible after six hours.

More recently, some type I families have been implicated in stress response. Two type I families, TisB-IstR and SymE-SymR, have LexA boxes upstream of the gene (Darfeuille *et al.* 2007; Kawano *et al.* 2007), indicating a role in the SOS response. When a cell undergoes SOS response, triggered by DNA damage, the LexA repressor is cleaved to allow transcription of downstream genes (Butala *et al.* 2009). For the TisB-IstR system, both the toxin and antitoxin could be deleted together or separately in their host *E. coli*- no cell death occurred (Vogel *et al.* 2004). But strains that are SOS-on, meaning no LexA repressor is present, could not be transduced with the toxin-only region unless antitoxin was provided in trans from a plasmid (Vogel *et al.* 2004). This suggests that the TisB toxin is indeed part of the SOS regulon.

During the stress response phase, translation decreases and protease activity increases. For type II systems, this results in less antitoxin being produced and more being degraded in the cytoplasm, (Christensen *et al.* 2004; Gerdes *et al.* 2005). The loss of antitoxin de-represses TA operon transcription and liberates toxin in the cytoplasm (Cataudella *et al.* 2012). But it is debated whether TAs have been selected for their benefits in inducing bacteriostasis or whether the effect is incidental. Any reduction in TA operon translation would be expected to liberate toxins, whether due to stress response or gene loss. It is quite plausible, then, that such toxin increase is secondary to the stress response and not an adaptive feature (Van Melderen and De Bast 2009; Goeders and Van Melderen 2014).

Furthermore, when five known type II TA systems were deleted from the *E. coli* chromosome there was no decrease in stress-fitness, as would be expected if the genes were an essential part of cells' adaptive response (Tsilibaris *et al.* 2007). The response

to amino acid starvation, nutrient limitation, and acidic environments were identical to wild-type, and they competed effectively with wild-type in co-culture (Tsilibaris *et al.* 2007). Similar results were seen with *Streptococcus mutans* deleted for a *relBE* and a *mazEF* locus, though mutants were more resistant to acid and less efficient in sugar usage in a biofilm (Lemos *et al.* 2005). There is always the possibility, then, that the remaining TA systems on the chromosome not deleted could be functioning in stress response, or that a particular locus simply hasn't been tested with the right stress to see a response. However, if that were the case then it begs the question of why the chromosome maintains so many apparently superfluous TA systems (Van Melderen 2010; Tsilibaris *et al.* 2007).

This has not been explored in the literature to date, but plasmid-borne TA systems should respond to stress conditions similarly to chromosomal TA systems. It may be that plasmids, which can conjugate into new cells even when host cells are no longer capable of replicating (Heinemann *et al.* 2000; Heinemann 1999; Willms *et al.* 2006), derive a selective advantage from inducing bacteriostatic states.

Persistence

TAs have also been implicated in persistence, an epigenetic trait that allows a small portion (10^{-6}) of a population to survive prolonged antibiotic exposure (Lewis 2005; Vazquez-Laslop *et al.* 2006). This is not a genetic mechanism: progeny are fully sensitive to antibiotics (Vazquez-Laslop *et al.* 2006). Persisters can exist in the population before exposure to antibiotics in response to random switching or environmental cues (Kint *et al.* 2012; Dörr *et al.* 2009; Dörr *et al.* 2010; Vega *et al.* 2012), or during antibiotic triggered SOS response (Shah *et al.* 2006).

Some individual loci have been strongly implicated in persistence. Deletion of the type I TA system TisB-IstR made cells unresponsive to ciproflaxin, which triggers SOS-mediated persistence (Dörr *et al.* 2010). Over-expression of the chromosomal type II TA toxin, HipA, in *E. coli* inhibits macromolecule synthesis, producing a dormant state in 95% of cells lasting three days (Korch and Hill 2006). A mutant allele, *hipA7*, confers high levels of persistence, about 10^3-10^5 more than non-mutant levels without overexpression (Korch *et al.* 2003). Interestingly, though, HipA7 is nontoxic (Korch and Hill 2006), suggesting that its ability to increase persistence might be independent of its affect on macromolecule synthesis.

But, while ectopic overexpression of TA operons like *hipAB* and *relBE* increases bacterial persistence (Korch and Hill 2006; Dörr *et al.* 2010; Vazquez-Laslop *et al.* 2006), so does overexpression of a number of other unrelated proteins, including ones that are non-toxic at normal expression levels (Vazquez-Laslop *et al.* 2006). This may reflect the heterogeneity seen in persistence mechanisms, with different mechanisms causing persistence in the presence of different stressors. A screen of 3,985 ORF knockouts in *E. coli* (out of 4,288) found that not a single strain failed to form persisters (Hansen *et al.* 2008), indicating there is no one gene necessary for the phenotype. Populations with reduced numbers of persister cells had mutations in global regulators like *hns*, which regulates the expression of a large number of genes in response to environmental signal, and genes like *hnr*, which regulates the sigma factor (RpoS) that initiates stationary phase (Hansen *et al.* 2008). Even the *tisAB* locus, strongly implicated in persister formation, only functioned during the exponential phase of growth and in response to ciproflaxin, not ampicillin or streptomycin (Dörr *et al.* 2010).

Like other proposed roles of TAs in stress response, there is evidence that the effect of TA systems as a group on persisters is incidental and not a selected advantage. Some bacterial persistence occurs in response to stress, while some bacteria that 'persist' were in a dormant state before the stressor occurred. The SOS response is stochastically activated in a fraction of exponential phase bacteria (Pennington and Rosenberg 2007). This has the associated effect of activating TA systems, as with stress response, further reducing translation rates (Fasani and Savageau 2013).

Because few deletions at individual TA loci have a significant effect on persistence (Dörr *et al.* 2010; Hansen *et al.* 2008; Vazquez-Laslop *et al.* 2006; Keren *et al.* 2004), it has been suggested that many TA loci could be working in aggregate to increase persistence (Dörr *et al.* 2010; Fasani and Savageau 2013). Persistence increases in *E. coli* as more type II loci are deleted, with a dramatic reduction in persistence when ten were deleted (Keren *et al.* 2004). But selection for individual operons would be weak. Thus it is unclear how many individual TA modules have a dedicated, adapted role in persistence. And as persistence can be activated before stress is administered, as with the HipA7 mutant (Korch *et al.* 2003), not in response to it, any such mechanism would have to be selected to deliberately keep a fraction of cells dormant during exponential growth.

Developmental hypothesis

Some of the strongest evidence for a physiological role of a TA system in a cellular context is seen with the type II mazF homologue in *Myxococcus xanthus* (*mazF-mx*), during fruiting body formation. Fruiting bodies formation is generally triggered under conditions of nutrient starvation, where cells coalesce to produce spores (myxospores) suspended on stalks (Berleman and Kirby 2009). During this process, 80% of cells undergo obligatory cell lysis and 20% become myxospores. Deletion mutants at the *mazF-mz* operon do not lyse, highlighting their importance for cell death under these condition (Nariya and Inouye 2008). Unlike most TA systems, *mazF-mx* is monocistronic, with the transcription factor MrpC taking the place of an antitoxin (Nariya and Inouye 2008). MrpC, a key developmental regulator, positively regulates *mazF-mx* transcription as part of a nutrient-sensitive kinase cascade, and inhibits toxin activity through direct binding (Nariya and Inouye 2008). This regulatory integration, and the strong deletion-associated phenotype, suggests that *mazF-mx* has been incorporated into host development.

Some roles proposed for chromosomal TAs still invoke replicon competition, as proposed for plasmid TAs. This is particularly true of anti-addiction, as well as genomic stabilization. By being refractory to deletion by homologous recombination, the TA would effectively be preventing incoming DNA from replacing it. Significant research has gone into identifying ways that chromosomal TAs are integrated into host networks. Specific loci, particularly *hipA7* in *E. coli* and *mazF-mx* in *M. xanthus*, have a well-characterized involvement with persistence and sporulation. These fall under the larger rubric of stress response, though attempts at categorizing all chromosomal TAs as general stress response loci has been weak as deletion of one to multiple loci in a strain rarely has a major effect on the cell (Van Melderen and De Bast 2009; Tsilibaris *et al.* 2007).

1.2.2 Restriction modification systems

Restriction modification (RM) systems involve an endonuclease function (restriction enzyme) and a methylation function. The endonuclease hydrolyses the phosphodiester backbone of DNA (cleavage) in a sequence dependent fashion. The methylase acts as an antitoxin by recognizing and modifying the same sequence as the restriction enzyme. Once methyl groups have been added to select nucleotides in a particular

sequence, it is no longer a substrate of the endonuclease. The interactions between these two functions have been shown to have PSK-like qualities. RM systems were first seen on chromosomes and thought to have a role in preventing phage infection, and later by extension all HGT, by cleaving incoming DNA unmethylated at the target sequence (Kobayashi 2001; Furuta *et al.* 2010).

There are four subtypes of RM systems. Type I RMs include three components, the restriction (R), the methylation (M) and the sequence recognition components (S). Together they assemble to form two multi-subunit enzymes, one for modification (SM) and one for restriction (SR) (Orlowski and Bujnicki 2008). The SR complex binds to unmodified recognition sequences in the DNA, and uses ATP to translocate the DNA towards itself. When the complex collides with another SR complex or a stalled replication fork, the DNA between them is cleaved (Studier and Bandyopadhyay 1988). At least some type I systems have been experimentally shown to not demonstrate PSK activity (O'Neill *et al.* 1997).

Type II systems involve two separately encoded enzymes, a restriction endonuclease and a methylase, each of which recognize the same palindromic sequence of DNA (Bujnicki 2003; Pingoud and Jeltsch 2001; Roberts *et al.* 2007). The restriction enzyme hydrolyzes phosphodiester bonds in the backbone of non-methylated DNA in the specific binding sites, a trait that has been extensively utilized in biotechnology. Expression of restriction enzymes can be delayed after host entry by a linked control gene, giving the methylase time to protect host DNA (Nakayama and Kobayashi 1998; Tao *et al.* 1991; Ives *et al.* 1992). Type II RM systems were shown to demonstrate PSK activity by Naito *et al.* (1995). Unlike the TA systems, which rely on differential lability of the toxin and the antitoxin, RM systems are believed to cause PSK primarily through dilution of the methylase: cell division after gene loss reduces the amount of methylase in the cell and increases the chance that a site in the chromosome will remain unmethylated, increasing its susceptibility to the restriction enzyme (Handa *et al.* 2000; Kobayashi 2001). Here, the lability resides in the interaction between the antitoxin and its substrate, as opposed to the interaction between the antitoxin and the toxin.

Type III systems have two components, R and M. The M component acts independently as a methylase, with both R and M in complex necessary for endonuclease activity (Orlowski and Bujnicki 2008; Sistla and Rao 2004; Furuta *et al.* 2010). The cleavage mechanism is unclear, but like Type I RM systems, the DNA is cleaved

through interaction with another complex on the same DNA molecule (Ishikawa *et al.* 2010; Raghavendra and Rao 2004). Type IV RM systems do not have their own methylase, but recognize and cleave sequences methylated by other RM systems (Bickle and Krüger 1993; Bourniquel and Bickle 2002). This could cleave methylated DNA of incoming viruses and plasmids, but may also trigger cell death when entering methylases start modifying the genome (Ishikawa *et al.* 2010; Furuta *et al.* 2010).

Distribution

RM genes have been found on plasmids, integrons, transposons and prophages as well as chromosomes (Furuta *et al.* 2010; Dempsey *et al.* 2005; Kita *et al.* 2003; Naito *et al.* 1995; Prakash-Cheng *et al.* 1993; Mruk and Kobayashi 2014). Chromosomal RM systems can be linked to recombination related genes such as integrases, invertases, and transposases (Furuta *et al.* 2010; Naito *et al.* 1995). There is evidence that the restriction enzyme itself may play a role in its movement and transfer between loci independent of mobile elements due to its ability to create double stranded breaks in genomic DNA, which could lead to shuffling of genomic regions during repair. Analysis of chromosomal RM loci between closely related species show an association between movement of a given locus and extensive genomic rearrangements (Furuta *et al.* 2010; Nobusato *et al.* 2000; Naderer *et al.* 2002).

The pattern of RM systems on chromosomes is similar to that seen with TA systems: both are ubiquitous on bacterial and archaeal chromosomes (Roberts *et al.* 2003; Nobusato *et al.* 2000; Furuta *et al.* 2010), with some species enriched and some, such as intracellular parasites, unusually poor in cut sites (Kobayashi 2001; Mruk and Kobayashi 2014). The presence of closely related loci found on phylogenetically distant organisms suggests HGT of all subtypes of RM systems (Nobusato *et al.* 2000; Furuta *et al.* 2010). Type II RM systems appear to be most prone to horizontal gene transfer, analogous to type II TA systems (Mruk and Kobayashi 2014), though type I and type III RM systems appear to undergo extensive domain swapping in the highly diverse target recognition domain (Naderer *et al.* 2002; O'Sullivan *et al.* 2000; Mruk and Kobayashi 2014).

Functions

As RM systems were discovered due to their effect on phage entry into the cell, they were originally cast as cellular defenses or immune systems. The effectiveness of RM systems against bacteriophages, though, is short lived (Kobayashi 1998) and more activities have been discovered since then, including PSK and chromosomal rearrangements (Naito *et al.* 1995; Ishikawa *et al.* 2010).

The evidence that RMs can exhibit PSK led to the proposal that they are primarily maintained as selfish elements, independent of their ability to aid the host (Naito *et al.* 1995; Kobayashi 1998). Their ability to exclude phage and incoming DNA becomes a form of replicon competition (Ambur *et al.* 2012), with competition between systems driving the exceptional diversity and specificity in RM sequence recognition (Kusano *et al.* 1995; Chinen *et al.* 2000). In a manner similar to anti-addiction of TA systems, an incoming RM cannot induce PSK in cells with a resident RM with the same sequence recognition (Kusano *et al.* 1995), as there would be a continuous supply of methylase to protect the cell even if the incoming methylase was lost. Competition within the host would select for RMs with less specific recognition sequences, and thus capable of cleaving more sites on the chromosome (Kusano *et al.* 1995; Chinen *et al.* 2000). RM systems display an extreme diversity of recognition sequences: if the primary purpose of RM systems was to cleave unmethylated foreign DNA entering the cell, such diversity would be unnecessary (Kobayashi 1998; Kusano *et al.* 1995; Chinen *et al.* 2000).

1.2.3 Bacteriocins

Bacteriocins are a group of bacteriocidal proteins with a relatively narrow killing spectrum, generally toxic to bacteria closely related to the producing strain (Riley and Wertz 2002). The toxin's mechanism of action and corresponding immunity function are diverse (Riley and Wertz 2002; Cascales *et al.* 2007). Almost all species of bacteria have been found to make at least one bacteriocin, either chromosomally or from plasmids (Riley and Wertz 2002; Klaenhammer 1988).

The best studied are the colicins, found on plasmids in *E. coli*. Toxins are either nucleases or pore forming proteins (Hardy *et al.* 1973). They are organized in an operon with an immunity function, and in most cases, a gene that encodes a protein that causes cell lysis. Expression of the lysis activity releases the toxin into the environment (Cascales *et al.* 2007). No colicin is produced under normal conditions,

though a small amount is always present (Cascales *et al.* 2007). After exposure to stress, toxin expression is induced in a small fraction of the cells (Mulec *et al.* 2003). It becomes the main product of the cell, released into the environment after lysis and killing neighboring cells lacking the immunity function. Though *E. coli* bacteriocins are exclusively plasmid borne (Riley and Wertz 2002), some closely related bacteriocins are not: pyocins of *Pseudomonas aeruginosa* and others are exclusively on the chromosomes, and bacteriocins of *Serratia marcesens* are found on both (Guasch *et al.* 1995; Duport *et al.* 1995; Riley and Wertz 2002).

Recently, Inglis *et al.* (2013) found that both a nuclease and a pore forming colicin was sufficient to stabilize a plasmid in monoculture over 180 generations, even under conditions where toxin expression was not strongly induced. Cells that lost the plasmid were no longer immune to the toxin, and thus susceptible to residual levels of colicin in the environment, creating a PSK phenotype on a population level (Inglis *et al.* 2013). This is similar to the action of antibiotics. Antibiotics are also toxins within the external environment, and can drive retention of antibiotic resistance mechanisms, essentially rendering an immunity function addictive. The linkage of the immunity and toxin function in bacteriocins allow them to promote their own spread and maintenance in bacteria.

1.2.4 ABI Mechanisms

Abortive infection (ABI) mechanisms are structurally and functionally diverse, united in their ability to trigger cell death upon infection with phage. They range from one to four genes, with toxins targeting membrane stability (Rex), tRNA, translation (AbiC), and phage DNA packaging (AbiE, AbiI, AbiQ) (Labrie *et al.* 2010; Parma *et al.* 1992). ABIs can be found on plasmids, chromosomes, intact and inactive prophages (Labrie and Moineau 2007) and phage (Parma *et al.* 1992). Of the over 20 described in *Lactococcus lactis* (Chopin *et al.* 2005), most are on plasmids.

Some ABIs are regulated at the transcriptional level. Toxin genes are tightly repressed, then released under conditions of phage infection, as with the *L. lactis* AbiD1 (Chopin *et al.* 2005). In other systems, the toxin is constitutively expressed as a dormant protein that becomes activated upon phage infection. The *E. coli* ABI Lit is directly activated by the Gol peptide, the major capsid protein of phage T4 (Bingham *et al.* 2000). The Rex system involves two proteins, RexA and RexB, co-expressed

constitutively from the coliphage λ lysogen (Parma *et al.* 1992). After infection, RexA is activated by a phage protein-DNA complex and in turn activates the RexB protein imbedded in the membrane (Labrie *et al.* 2010). RexB drains membrane potential, halting ATP production and affecting both cell multiplication and aborting phage infection (Parma *et al.* 1992). The exact mechanism of action of ABI systems is not always known, even for systems such as *pif* on the F plasmid, which was one of the first ABIs discovered (Condit 1975; Cram *et al.* 1984). Exclusion of phage T7 by the *pif* operon (the exclusion protein PifA and its autoregulatory repressor PifC) is characterized by membrane leakage and inhibition of macromolecular synthesis (Cram *et al.* 1984; Cheng *et al.* 2004). Only recently, the PifA protein has been shown to localize to the cytoplasmic membrane, and has been suggested to act by recruiting T7 proteins and inhibiting essential membrane proteins (Cheng *et al.* 2004).

There is crossover between the ABI and PSK phenotype. Many genes that act as ABIs are also able to act as PSKs: some, but not all, ABIs induce PSK (Dy *et al.* 2014) and many TA and RM systems can act as ABIs. The RnlAB system of *E. coli* reacts to T4 infection, which activates RnlA, an endoribonuclease that degrades late phage mRNAs (Koga *et al.* 2011). AbiQ was first identified as an ABI module then shown to be a type III TA system (Samson *et al.* 2013). Other TA systems shown to induce ABI include MazEF (against P1) (Hazan and Engelberg-Kulka 2004), Hok-Sok (against T4) (Pecota and Wood 1996), type III system ToxIN (against A2 and M1) (Blower *et al.* 2011; Fineran *et al.* 2009) and type IV system AbiE (Dy *et al.* 2014). On the other hand, of 12 known two-component ABI systems, only three contained a protein toxic upon over-expression in *E. coli* and of those only one, AbiE, also contained an antitoxin function (Dy *et al.* 2014). AbiE was further shown to function as a type IV TA system, and stabilize plasmids (Dy *et al.* 2014).

Phage infection perturbs macromolecular synthesis (Pecota and Wood 1996), which can activate resident TA systems, causing cell death without gene loss- a form of ABI. Resident RM systems can also induce cell death when the host is infected by RMs with the same control gene (Nakayama and Kobayashi 1998). Many type II RMs have a control gene that delays expression of the endonuclease upon entry into the host (Mruk and Blumenthal 2008). A resident RM with the same control gene can force early expression of the incoming RM system, resulting in cleavage of host DNA before the incoming methylase can protect it (Nakayama and Kobayashi 1998; Chinen *et al.* 2000).

Some ABI systems can be activated by things other than phage infection, which suggests that they may have additional roles in cell physiology. Though the toxins are effective at halting phage infection, they also cause cell death and at high levels can do so independent of phage infection (AbiK, AbiB) (Chopin *et al.* 2005). Cheng *et al.* (2004) suggests that the specificity of some ABIs like PifA makes it unlikely to function only in phage exclusion: resistance to the target phage, T7, would only be expected to select for insensitive phages such as T3. They speculate that the membrane-bound protein has additional functions in the biology of F plasmids, possibly monitoring triphospate pools (Cheng *et al.* 2004). The *rex* system has been shown to inhibit proteases that activate during nutrient stress, and has been proposed to function during stationary phase in uninfected cells. (Engelberg–Kulka *et al.* 1998; Slavcev and Hayes 2003).

1.3 Exploring the parameter space of PSK

All PSKs identified thus far include a toxin and an antitoxin/immunity function. There are many such genes in the bacterial gene pool, but only some types induce PSK, including the canonical TA systems, RM systems, bacteriocins, and ABIs (Figure 1.2). Even within these types of genetic modules, there is diversity amongst families and individual loci in their exhibition of PSK. What defines the boundaries of the overlap between the biochemical functions of toxin and antitoxin and the phenotype of addiction?

In this thesis, I investigate what differentiates TA or putative TA systems that produce a PSK phenotype from systems that do not. I address this in two parts, here and in the thesis as a whole. First, I look at differences within known or predicted families of TA systems. Using bioinformatics, I test the apparently narrow phylogenetic distribution of type I TA systems, and the relationship this may have to gene mobility and PSK. I also look experimentally at a specific family of genes proposed to be a type I TA system, PT-ptaRNA1 that exhibits a structure and distribution similar to TAs that induce PSK. Second, I look at a pair of genes taken from the general pool of toxin and antitoxin gene pairs, and explore what is necessary for these genes to be active PSK systems.

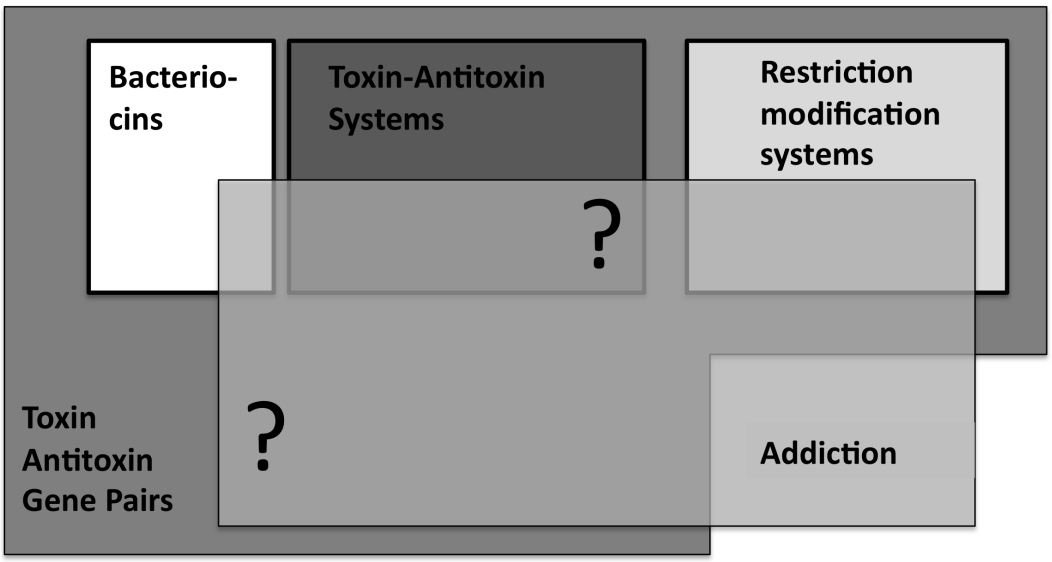


Figure 1.2: The intersection between subgroups of toxin and antitoxin gene pairs, and the addiction phenotype. The features that determine whether a given system is addictive have not been fully defined. Some types of toxin and antitoxin gene pairs, like bacteriocins and TA systems, are addictive while some are not. Even within these types, only some individual loci exhibit the phenotype.

1.3.1 Part I: Distribution and functionality of computationally identified PSKs

Numerous surveys of bacterial genomes have been performed to determine the distribution of TA systems. The distribution of most type II families across replicons and across phyla is generally considered to be more indicative of HGT than that seen with many type I families (Leplae *et al.* 2011; Fozo *et al.* 2010; Mruk and Kobayashi 2014; Makarova *et al.* 2009). Location on mobile elements allows genes to move between cells (Thomas and Nielsen 2005). Some mobile elements that have integrated into chromosomes will excise later, potentially taking chromosomal genes with them, while others can lose their mobility and be primarily vertically inherited (Ochman *et al.* 2000; Sundin 2007). For example, a 148-kb plasmid in *Pseudomonas syringae* was found capable of replicating autonomously and integrating into the chromosome via insertion sequence elements (Sundin 2007). The plasmid excised from the chromosome into smaller plasmids, leaving parts of the original plasmid in the chromosome and taking as much as 47-kb of flanking chromosome with it.

This type of movement leaves some previously mobile genes on the chromosome, mobilizes previously chromosomal genes, and increases the sequence similarity of the excised plasmid to the chromosome, thus increasing the probability of subsequent integration events. Genes that are frequently mobile across a wide range of species can exhibit a distribution that is independent of the host phylogeny (lineage independent) (Ochman *et al.* 2000; Sundin 2007). This is the pattern seen with type II TA systems, but not type I.

Not only is the distribution of type II TAs on chromosomes lineage independent, chromosomal TA loci are often on genomic islands, which is taken as evidence of recent HGT (Ramage *et al.* 2009). Phylogenetic analysis of the type II toxin MazF suggested frequent interchange between plasmids and chromosomes (Chopra *et al.* 2013). Homologues of the toxin clustered into five groups, with plasmid-borne members interspersed with chromosomal clusters (Chopra *et al.* 2013). Apparent gene transfer events were particularly common with toxins associated with the AbrB group of antitoxins. Antitoxins with AbrB DNA binding domains (Vaughn *et al.* 2000), found associated with toxins from the MazF, CcdB, and VapC families (Leplae *et al.* 2011), are often flanked by transposon like elements (Gerdes *et al.* 2005) which could account for their wide distribution.

While there is extensive evidence for the movement of type II TA systems between chromosomes and plasmids, many type I systems have so far only been found on chromosomes. Only the families Hok-Sok, Fst-RNAI, and the putative TA system plasmid_Toxin-ptaRNA1 (PT-ptaRNA1)have been identified on plasmids (Findeiß *et al.* 2010; Gerdes *et al.* 1986; Gerdes *et al.* 1986; Weaver *et al.* 2009). Type I systems are also more likely to be found within a single clade of bacteria, rather than dispersed across phyla. Most of the chromosomal-only systems have not been tested for their ability to demonstrate PSK, those that have did not exhibit it (Kawano *et al.* 2007; Fozo *et al.* 2008). The reason for these differences between type I and type II families has not been elucidated, though similar differences can be seen in the distribution of type II RM systems, at least some of which exhibit PSK, and type I RM systems, which are not yet known to (Mruk and Kobayashi 2014).

Much of the known distribution of TA systems relies on computational analysis of sequence data. Historically, type I systems have been difficult to detect in silica, in part due to the small size of the toxin (under 60 amino acids) and their RNA antitoxin. The known range of type I systems across replicons and phyla is in flux, as better methods to detect loci are rapidly developed and new systems are found. A recent bioinformatic search done by Fozo *et al.* (2010) was able to greatly expand the number of known toxin loci across bacterial lineages, and a number of new families have been described computationally and experimentally (Fozo *et al.* 2010; Findeiß *et al.* 2010; Kawano 2012; Sayed *et al.* 2012; Fozo *et al.* 2008).

I applied bioinformatic methods to investigate the distribution of different families of type I TA systems, in an attempt to broaden the known range of supposedly narrow type I TAs. Novel type I systems were searched for, as well as novel loci from known families. Comparisons were made between the phylogenetic range of known type I families, from data here and elswhere, and known type II TA systems, and the degree to which they are found on mobile elements.

Whether the distribution of a TA system is reflected in an ability to exhibit PSK and engage in a horizontal lifestyle was tested experimentally with a proposed family of type I TA system that has a distribution more commonly seen in type II systems (Findeiß *et al.* 2010; Fozo *et al.* 2010). The PT-ptaRNA1 family has primarily been described computationally (Findeiß *et al.* 2010). Interestingly, the distribution of the operon (PT-ptaRNA1) is lineage independent, occurring across diverse species (Findeiß *et al.* 2010). The authors suggested that the distribution was a result of

widespread HGT. I tested the system experimentally to see if it exhibited biochemical traits of a TA system, including toxicity of the toxin and exhibition of PSK.

1.3.2 Part II: Ability of a given toxin and antitoxin gene pair to induce PSK

In the second part of the thesis, I look at the difference between toxin and antitoxin gene pairs that create a PSK phenotype and those that do not. As a model system, I used the toxin and antitoxin gene pair barnase and barstar from *Bacillus amyloliquefaciens*, which can be found unlinked on the chromosome. Barnase is a small (12.4kDA) RNase (Hartley and Smeaton 1973), bound with high affinity (kd 10^{-14}M) by barstar, its intracellular inhibitor (Wang *et al.* 2004; Hartley 1993). Together, they are not known to exhibit PSK: I am interested in what conditions are necessary for them to do so.

Functions of barnase and barstar

The physiological or ecological functions of barstar and barnase are unknown, though they are often assumed to help the cell to acquire nutrients from the environment. In support of this, *Bacillus* spp. are adapted to low nutrient habitats (soil), and secrete a host of extracellular degradative enzymes for nutrient uptake prior to sporulation (Kharitonova and Vershinina 2009). *Bacillus* spp. have pathways for importing extracellular ribonucleic acids (Saxild and Nygaard 1991; Beaman *et al.* 1983) and either re-incorporating them into their RNA (Bodmer and Grether 1965) or catabolising them as a source of nutrients (Schuch *et al.* 1999). The biosynthesis of a variety of related guanyl-specific RNases can be induced by phosphorous (Znamenskaya *et al.* 1995; Ulyanova *et al.* 2011) and nitrogen limitation (Kharitonova and Vershinina 2009). But the evidence for barnase to act similarly is inconstistent. The barnase gene is flanked by genes involved in nitrogen, phosphorous and carbon metabolism (Ulyanova *et al.* 2011). Yet the expression of barnase does not appear to be regulated by low levels of phosphorous and nitrogen in the environment (Znamenskaya *et al.* 1995; Ulyanova *et al.* 2011). This would be expected if its primary role was to raise nutrient levels in a manner similar to related RNases.

Other theories on the role of barnase have considered the possibility of barnase acting as a toxin on neighboring cells, instead of or in addition to a role in nutrient scavenging. For example, Ramos *et al.* (2006) created a heterologous expression system in *E. coli* and found that barnase expression could lead to zones of inhibition in surrounding lawns composed of other *E. coli* or other soil bacteria. The presence of barstar in the test strain used for the lawn eliminated the zone of inhibition. The authors suggested that barnase could act similarly to bacteriocins (Kleanthous 2010).

The two views on the role of barnase are not necessarily exclusive. It has also been speculated, for example, that barnase could aid in nutrient uptake at the onset of sporulation by killing neighboring cells (Ulyanova *et al.* 2011), a process known as cannibalism. Two such systems have been described in *B. subtilis*, *skf* and *sdp*, under control of the Spo0A transcription factor (González-Pastor 2011; Claverys *et al.* 2006). As *B. amyloliquefaciens* lacks known cannibal operons, barnase could fulfill this function. Similar to cannibal operons, barnase is not expressed in null-Spo0A mutants and has been reported as stationary phase dependent (Paddon *et al.* 1989). Barstar may have a binding site for the SigW transcription factor, which regulates cell immunity to toxins, and is inactive in null-Spo0A bacteria (Ulyanova *et al.* 2011). As such barstar expression may also be SpoA regulated. Most work done on barnase expression and function has been preliminary and isolated. Further research will be necessary before strong statements can be made about the role of this protein in *B. amyloliquefaciens* in its natural environment.

Biochemistry and secretion of barstar and barnase

Biochemically, barnase is guanyl-specific, cleaving single stranded RNA downstream of guanine nucleotides with a preference for downstream purines (Mossakowska *et al.* 1989; Day *et al.* 1992; Bastyns *et al.* 1994). The RNA is cleaved via a two-step process (Figure 1.3), as diagrammed below. The activity of barnase inside the cell is inhibited by a one to one binding to barstar (Hartley and Smeaton 1973). The negatively charged amino acids on barstar sterically block the positively charged active site on barnase through strong electrostatic binding (Buckle and Fersht 1994) (Figure 1.4). This interaction is extremely strong, with a dissociation constant between the two proteins at 10^{-14} M. The dissociation constant is the concentration of protein at which

half is unbound to the other protein, such that the more tightly the two interact the lower the dissociation constant will be.

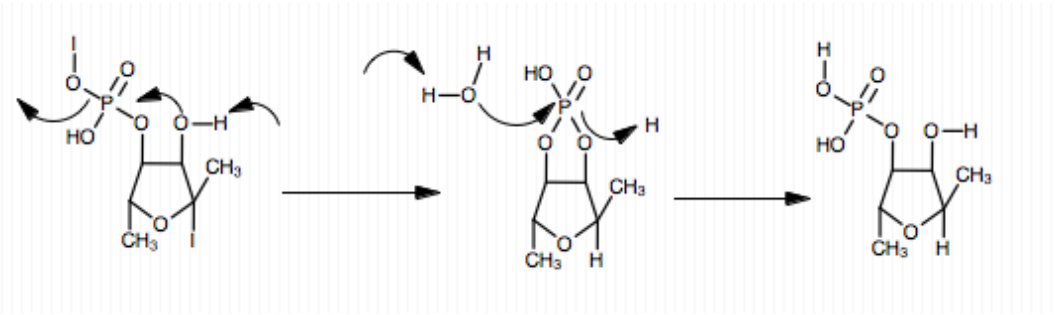


Figure 1.3: Hydrolysis of RNA by barnase. Barnase is a guanyl-specific RNase. Single stranded RNA is cleaved downstream of guanine nucleotides in a two-step process, with a cyclic phosphate intermediate formed (the transesterification step) followed by hydrolysis of the intermediate, yielding a new guanosine 3′ phosphate and 5′-OH end (Rushizky *et al.* 1963). The general base is the amino acid glu73, while the general acid is his102 (Schreiber *et al.* 1997; Mossakowska *et al.* 1989).

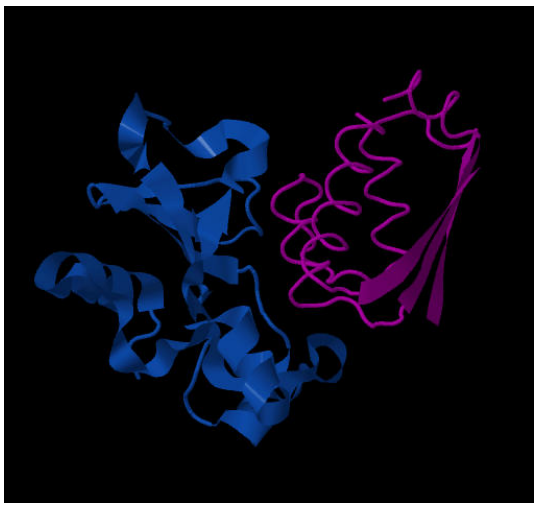


Figure 1.4: Interactions between barnase and barstar. Barstar (pink) has negatively charged amino acids that interact with and sterically block the active site of barnase (blue). Conformation from (Buckle and Fersht 1994), as viewed in Jmol: an open-source Java viewer for chemical structures in 3D. `http://www.jmol.org/`

In its native host, *B. amyloliquefaciens*, barnase is translated as a pre-proprotein (Paddon *et al.* 1989). The 26 amino acid pre sequence is an export signal that is cleaved during translocation of the cell membrane (Chen and Nagarajan 1993). Export of barnase on its native signal sequence is slow relative to other signal sequences: barnase has been shown in the cytoplasm up to 90 minutes after translation in *B. subtilis*. This may account for the presence of barstar (Chen and Nagarajan 1993), as the analogous but rapidly exported RNase binase, from *Bacillus intermedius*, apparently lacks an intracellular inhibitor (Ulyanova *et al.* 2011). Barnase is exported independently of barstar, allowing it to be active extracellularly (Paddon *et al.* 1989).

The pro-sequence (13 amino acids) is cleaved by extracellular serine proteases after secretion, to make the mature barnase protein (Paddon *et al.* 1989). Its presence does not appear to affect overall refolding rates or catalytic activity of barnase (Gray *et al.* 1993). It does increase protein binding to chaperon GroEL, potentially aiding in transport of the protein (Gray *et al.* 1993; Chen and Nagarajan 1993). Barnase is also bound by the *E. coli* chaperon SecB (Stenberg and Fersht 1997). SecB transfers proteins to a translocase that secretes proteins through the cell membranes and is relatively conserved between *E. coli* and *B. subtilis* (Harwood and Cranenburgh 2008).

Barstar and barnase as a model system

The biochemical activity of barnase and barstar is similar to that seen in type II TA systems. The type II toxins MazF, HigB and VapC are all ribonucleases (Van Melderen and De Bast 2009), also tightly bound by intracellular inhibitors (Brzozowska and Zielenkiewicz 2013). As a secreted toxin, barnase also has some similarities to bacteriocins. These similarities, plus the well-studied biochemistry of the proteins, prompted their selection to test the necessary conditions for a given toxin and antitoxin to become a PSK.

There is precedent for genes to become addictive in new environmental and genetic contexts. As genes move from locations between and within replicons and between different cells and external environments, both their function and the selective pressures acting on them can change (Heinemann and Silby 2003). Context, for example,

affects the spread of antibiotic resistance via PSK-like mechanisms.

Genetic context changes as genes move from replicon to replicon. Many antibiotic resistance genes were originally chromosomal genes, being selected for as resistance mechanisms on plasmids (Martinez 2012; Baquero *et al.* 2008). Environmental context changes as genes move into new cells in new environments. Some resistance genes now seen on plasmids originated from producer strains, to protect the bacterium from its own antibiotics, or strains that co-exist with producer strains (Laskaris *et al.* 2010; Benveniste and Davies 1973; Davies and Davies 2010). But some genes that function as antibiotic resistance mechanisms have alternate functions in their original host (Davies and Davies 2010; Forsberg *et al.* 2012; Martinez 2012; Laskaris *et al.* 2010). The gene *qnrA*, which encodes resistance to quinolines, comes from the non-antibiotic producing aquatic bacterium *Shewanella algae* (Poirel *et al.* 2005). Now it has become an important pathway for resistance to synthetic antibiotics such as quinolines.

Once the genes become mobilized, they can potentially spread to new hosts. Lacking their original biochemical and genetic context, they can still be maintained and selected for if they confer antibiotic resistance, a form of exaptation (Baquero *et al.* 2008). In the presence of lethal concentrations of antibiotics in the environments, these resistance genes become addictive and spread quickly through populations (Cooper and Heinemann 2000).

I am interested in using barstar and barnase to test the importance of cellular and genetic context for expression of the PSK phenotype. Within the cell, I consider the relative expression levels and stability of the different components necessary for barstar and barnase to exhibit PSK in a manner similar to type II TA systems. The role of genetic context is explored by testing the hypothesis that a given secreted toxin can induce PSK when expressed from a plasmid, in a manner analogous to antibiotics.

Ultimately, I am interested in the evolution of PSK systems in bacteria. PSKs have successfully colonized large portions of the bacterial gene pool, proliferating on both mobile elements and chromosomes. The movement and interactions of genes in new contexts has the potential both to drive evolution of PSKs and select for their maintenance. The ability of genes to become addictive in new contexts, and selection for such phenotypes on mobile elements through competition, has implications for selection of genes onto plasmids and subsequent spread through bacterial populations.

Chapter 2

Materials and Methods

2.1 Bacterial Strains and Culture Conditions

Escherichia coli (*E. coli*) (Table 2.1) was cultured in Luria-Bertani broth (LB; Invitrogen), M9, or RM media, supplemented with antibiotics (AB) and amino acids as necessary (Appendix A). Bacteria were cultured in a dry incubator (Sanyo) on petri dishes or in liquid broth in at Gyrotory Water Bath Shaker (New Brunswick Scientific Co. Inc.) at 250 rpm for aeration, unless otherwise stated. *E. coli* was primarily cultured at 37℃: strains containing temperature-sensitive constructs were grown at 30℃. Master stocks were stored at -80℃ in 15% glycerol.

Table 2.1: Strain genotypes

Strain Designation[1]	Genotype	Source
CSH100 (CS)	F', ara-600, Δ(gpt-lac)5, λ-, relA1, spoT1, thiE1 F128-(CSH100) lacI^q, lacZ	(Miller 1992)
CSH10 (C4)	F', ara-600, Δ(gpt-lac)5, λ-, relA1, spoT1, thiE1 F128-(CSH104) lacI, lacZ	(Cupples and Miller 1989)
LMG194 (LM)	F-ΔlacX74galE thi rpsL ΔphoA (pvuII) Δara714 leu::Tn10	(Invitrogen)
Rel606 (R6)	F-, tsx-467(Am), araA230, lon-, rpsL227(strR), hsdR-, [mal+](LamS) (an *E. coli* B strain)	(Jeong2009)
MG1655 (MG)	F-, λ-, rph-1	(Guyer *et al.* 1981)
DH5α (DH)	F-, Δ(argF-lac)169, ϕ80d*lacZ*58(M15), ΔphoA8, glnV44(AS), λ-, deoR481, gyrA96(NalR), recA1, endA1, thiE1, hsdR17 (R_K- M_K+)	(Grant *et al.* 1990)

Table 2.1: Strain genotypes

Strain Designation[1]	Genotype	Source
XL1 Blue (XL)	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ $lacI^Q$ Δ(lacZ)M15] hsdR17(R_K- M_K+)	(Stratagene)
BL21 (DE3) (BL)	F- ompT hsdSB(R_B-, M_B-) gal dcm (DE3)	(Studier and Bandyopadhyay 1988; Jeong *et al.* 2009)

1 In parenthesis, the two-letter designation used to denote parent of derived strains. i.e. CSH100 with plasmid pBB02 designated CSBB02, etc.

2.2 Cloning

$CaCl_2$ transformation

Bacteria were made competent by growing to an optical density of 0.2-0.4 at 600 nm (OD_{600}), then pelleted by centrifugation at 13.5 g and resuspended in 0.1 M cold $CaCl_2$. Bacteria were then incubated on ice for 2-4 hrs, pelleted at 5 krpm and resusupended in 0.85 M cold $CaCl_2$ with 15% glycerol for a total volume 1/30 of initial culture. Bacteria were used immediately or frozen at -80℃. For transformation, ~50 ng of DNA was added to 100 μl of competent cells and incubated for 30 min. Bacteria were heat shocked at 42℃ for 90 sec then placed on ice for 5 min before inoculating into LB for culture for 40 min at 37℃ or 90 min at 30℃. Some ampicillin (AMP) resistant transformants were plated directly. Both strains XL1 Blue and DH5α were used for routine cloning.

Polymerase chain reaction (PCR)

Standard PCR was performed with Taq DNA polymerase (Bioline), 1x buffer, 10 μM dNTPs, 50 μM MgCl3, and 10 μM primer (each). Unless stated otherwise, the PCR program was 30 cycles involving a denaturing step at 94℃ and an extention step at 72℃. Proofreading PCR was performed with Phusion High-Fidelity polymerase (ThermoScientific), according to directions.

Isolating, cutting and annealing DNA

Plasmid preparations were carried out with the PureLink®Quick Plasmid Miniprep Kit (Invitrogen), and PCR and gel clean-ups with Wizard®SV Gel and PCR Clean-Up System (Promega). Nucleic acids were quantified with a ND-1000 Spectrophotometer nanodrop (Biolab). Ligations were typically performed with T4 ligase (ThermoScientific), with 60 ng vector and 3:1 insert to vector ratio (by kilobase (kb)). Additional DNA modifications were made with shrimp alkaline phosphatase (ThermoScientific). Restriction enzymes (RE) were purchased from Roche and Fermentas, and used according to manufacturer's specifications. Routine cloning kits used include CloneJET Cloning kit (ThermoScientific), pGEM-T Easy Vector System (Promega). The kit pBAD Topo TA Expression (Invitrogen) was used to place genes under a P_{BAD} promoter, and pSMART GC-LK cloning and amplification kit (Lucigen) was used to place genes on a pBR322 backbone. Individual cloning steps as described below (Table 2.2) using specified primers (Table 2.3).

Table 2.2: Plasmids and cloning

Plasmid	Genotype and cloning
pBB01	P_{tac}-*phoA-bar* P_{con}-*yrdF*, pUC19 backbone. High copy number. *bla*
	Plasmid contains genes for barnase (*bar*) under the *phoA* signal sequence and barstar (*yrdF*), and AMP resistance (*bla*). Gift from Dr. Thomas Fox. Previously pMT416. Described (Hartley 1988)
pBAD Topo TA	*araC* P_{BAD}- EK-(TA)-V5-His, pUC19 backbone. High copy number. *bla*
	Vector from pBAD Topo TA expression kit, with the AraC control gene (*araC*), an enterokinase recognition site (EK), a TA insertion site downstream the *ara* induced promoter (P_{BAD}), a V5 epitope (V5) and a polyhistidine region (His). Invitrogen.
pBADLacZ	P_{BAD}-*lacZ*-V5-His (prevously pBAD-TOPO/*lacZ*/V5-His), pUC19 backbone. High copy number. *bla*
	Control plasmid with inducible expression of β-galactosidase fragment (*lacZ*), provided in pBAD Topo TA Expression kit. Invitrogen.
pBB02	P_{BAD}-*phoA-bar* P_{con}- *yrdF*, pUC19 backbone. High copy number. *bla*
	The fragment <*phoA-bar* P_{con}- *yrdF* >was amplified from pBB01 using primers B:BNcoI_F and B:BPmeI_R (Table 2.3). The vector pBAD Topo TA was cut with REs PmeI and NcoI to remove EK recognition site, V5 epitope and His tag, and ligated with similarly cut insert.
pBS01	P_{tac}- *yrdF*, pIDT-SMART backbone. High copy number. *bla*

Table 2.2: Plasmids and cloning continued ...

Plasmid	Genotype and cloning
	Sequence synthesized by Integrated DNA technologies, (Appendix B).
pBB05	P_{tac}- *yrdF ara* C P_{BAD}- *phoA- bar* P_{con}- *yrdF*, pUC19 backbone. High copy number. *bla*
	The fragment <*ara* C P_{BAD}- *phoA- bar* P_{con}- *yrdF* >was amplified from pBB01 using primers araCBAD_BB F and araCBAD_BB R (Table 2.3) and inserted into pJET via the CloneJet PCR cloning kit, giving pJET11. The fragment <P_{tac}- *yrdF* >was excised from pBS01 with NotI and XhoI, and inserted into corresponding sites on pJET11, deriving pBB05
pBB06	P_{tac}- *yrdF araC* P_{BAD}- *phoA- bar*, pUC19 backbone. High copy number. *bla*
	The fragment <P_{tac}- *yrdF ara* C P_{BAD}- *phoA- bar* >was amplified from pBB05 using primers bngene_R and T7_F (Table 2.3). The fragment was cloned into pGEM T-Easy cloning vector.
pBB10	P_{tac}- *yrdF araC* P_{BAD}- *bar* P_{con}- *yrdF*, pUC19 backbone. High copy number. *bla*
	The fragment <*bar* P_{con}- *yrdF* >was amplified from pBB02 using primers Bn_ncoI_F and Barstar_ncoI_R (Table 2.3). The fragment was cut with REs NcoI, and ligated into pBB05, also cut with NcoI to remove the <*phoA- bar* P_{con}- *yrdF* >fragment.
pSMART GC LK	GC cloning site, pBR322+Rop backbone. Mid copy number (~20). *nptII*
	pSMART GC LK cloning kit, conferring resistance (*nptII*) to kanamycin (KAN). Lucigen.
pLKCat	*cat*, pBR322+Rop backbone. *nptII*, *cat*
	Phosphorylated M13 (Table 2.3) primers were used to amplify chloramphenicol (CM) resistance gene (*cat*), originally from pBSL181 (Alexeyev and Shokolenko 1995) and inserted into pSMART GC LK cloning kit.
pLKMCS	Multiple cloning site (MCS), pBR322+Rop backbone. *nptII*.
	MCS was amplified from pBluescript KS using phosphorylated primers blsc_MCS_F and blsc_MCS_R (Table 2.3), and inserted into pSMART GC LK cloning kit.
pLK_Par	*parDE*, pBR322+Rop backbone. Mid copy number. *nptII*
	Contains the ParDE type II TA system (*parDE*). Gift from Ryan Catchpole
pBRA	pBR322+Rop. Mid copy number. *bla*
	pBR322 with the tetracycline resistance gene (*tetA*) removed, gift from Dr. Brigitta Kurenbach

Table 2.2: Plasmids and cloning continued ...

Plasmid	Genotype and cloning
pBRT	pBR322+Rop. Mid copy number. *tetA*
	pBR322 with *bla* removed, gift from Dr. Brigitta Kurenbach
pBB03	P_{tac}-*phoA*- *bar* P_{con}- *yrdF*, pBR322+Rop backbone. *nptII*
	The fragment <P_{tac}-*phoA*- *bar* P_{con}- *yrdF* >were amplified from pBB01 using phosphorylated primers BBpmeI_R and P_{tac}:barnase_F and inserted into pSMART GC LK cloning kit.
pBS03	P_{tac}- *yrdF* (Table 2.3), pBR322+Rop backbone. *nptII*
	The fragment <P_{tac}- *yrdF* >was amplified from pBS01 using phosphorylated primers P_{tac}:bs_F and P_{tac}:bs_R (Table 2.3), and inserted into pSMART GC LK cloning kit.
pHSG415	Temperature sensitive, pSC101 derived replicon. Low copy number (4-5). *bla*, *cat*, *nptII*
	Gift from Dr. Kobayashi. Described (Hashimoto-Gotoh *et al.* 1981)
pTN9	*paeR7*, pHSG415 backbone. Low copy number. *cat*, *nptII*, *bla*
	Contains the PaeR7 RM system (*paeR7*). Gift from Dr. Kobayashi. Described (Naito *et al.* 1995)
pIK172	*ecoRI*, pHSG415 backbone. Low copy number. *cat*
	Contains the EcoRI RM system (*ecoRI*). Gift from Dr. Kobayashi. Described (Naito *et al.* 1995)
pHS_Par	*parDE*, pHSG415 backbone. Low copy number. *cat*, *nptII*, *bla*
	Gift from Ryan Catchpole.
pHS_PT	Series of plasmids with ptaRNA1 operons on temperature sensitive vectors.
pHS_PA	*pr1_PA* pHG415 backbone. Low copy number. *bla*, *cat*, *nptII*
	Putative PT_PA- ptaRNA1 operon from *Pseudomonas aeruginosa* plasmid pMATVIM-7 (*pr1_PA*, locus tag ECOK1357_5213) was synthesized as in (Appendix B) and inserted into the BamHI site of pHSG415 by GenScript. An extra 150-200 bps 5' of the gene was synthesized to account for the long 5' UTRs found in some type I systems. Part of the HS_PT series.
pHS_EC	*pr1_EC*, pHG415 backbone. Low copy number. *bla*, *cat*, *nptII*
	Putative ptaRNA1 loci from *E. coli* OK1357 (*pr1_EC*, locus tag AM778842.1) was synthesized as in (Appendix B) and inserted into the BamHI site of pHSG415 by GenScript. An extra 150-200 bps 5' of the gene was synthesized to account for the long 5' UTRs found in some Type I systems. Part of the HS_PT series.

Table 2.2: Plasmids and cloning continued ...

Plasmid	Genotype and cloning
pHS_BB	P_{tac}-*yrdF ara* C P_{BAD}- *phoA- bar* P_{con}-*yrdF*, pHSG415 backbone. Low copy number. *bla*, *nptII*
	Synthesized and cloned by GenScript into the *cat* gene of pHSG415. Sequence as (Appendix B).
pBAD33	P_{BAD}-MCS, pBR322+ROP backbone. Mid copy number (~20). *cat*
	Gift from Dr. André O. Hudson. (Guzman *et al.* 1995)
pBD_PT	Series of plasmids with plasmid_Toxins on pBAD33 for expression under P_{BAD}
pBD_AB	P_{BAD}-*pt_AB*, pBR322+ROP backbone. Mid copy number. *cat*
	Putative plasmid toxin (PT) ORF from *Acinetobacter baumannii* ATCC 17978 (*pt_AB*, locus tag A1S_0001) was synthesized with a T7 ribosomal binding site as in (Appendix B) and inserted into the SacI site of pBAD33 by GenScript. Part of the BD_PT series.
pBD_XC	P_{BAD}-*pt_XC*, pBR322+ROP backbone. Mid copy number. *cat*
	Putative PT ORF from *Xanthomonas campestris* pv. vesicatoria strain (*pt_XC*, locus tag XCV2162) was synthesized with a T7 ribosomal binding site as in (Appendix B) and inserted into the SacI site of pBAD33 by GenScript. Part of the BD_PT series.
pBD_BP	P_{BAD}-*pt_BP*, pBR322+ROP backbone. Mid copy number. *cat*
	Putative PT ORF from *Burkholderia pseudomallei* strain K96243 (*pt_BP*, locus tag BPSL3261) was synthesized with a T7 ribosomal binding site as in (Appendix B) and inserted into the SacI site of pBAD33 by GenScript. Part of the BD_PT series.
pBD_EC	P_{BAD}-*pt_EC*, pBR322+ROP backbone. Mid copy number. *cat*
	Putative PT ORF from *E. coli* OK1357 (*pt_EC*) was synthesized as a BioBlock (Integrated DNA Technologies, (Appendix B)) as RBS:pl_tx_E.coli with a T7 ribosomal binding site, and amplified with M13 primers (Table 2.3). The amplicon was cut with SacI and inserted into the SacI site of pBAD33. Part of the BD_PT series.
pBD_PA	P_{BAD}-*pt_PA*, pBR322+ROP backbone. Mid copy number. *cat*
	Putative PT ORF from *P. aeruginosa* plasmid pMATVIM-7 (*pt_PA*) was amplified from pHS_PA using primers RBSpl_tox_PA_F and pl_tx_PA_R (Table 2.3), adding a T7 ribosomal binding site. The amplicon was cut with SacI and inserted into the SacI site of pBAD33. Part of the BD_PT series.
pGEM-3Z	P_{Lac}-*lacZ* (MCS)-P_{T7}, pUC19 backbone. High copy number. *bla*
	Supplied by Promega as a control plasmid.

Table 2.2: Plasmids and cloning continued ...

Plasmid	Genotype and cloning
pGEM T-Easy	P_{Lac}-*lacZ* (MCS)-P_{T7}, pUC19 backbone. High copy number. *bla*
	Promega TA cloning vector for cloning PCR products.
pGEM_PT	Series of plasmids with PTs on pGEM T-Easy vectors.
pGEM_ECT	P_{T7}-*pt_EC*, pGEM T-Easy vector. High copy number. *bla*
	Fragment *pt_EC* was synthesized as a BioBlock (Integrated DNA Technologies, (Appendix B)) as RBS:pl_tx_E.coli with a T7 ribosomal binding site, and amplified with M13 primers (Table 2.3). The amplicon was inserted into the pGEM T-Easy vector and sequenced to confirm construct orientation downstream of the T7 promoter. Part of the pGEM_PT series.
pGEM_PAT	P_{T7}-*pt_PA*, pGEM T-Easy vector. High copy number. *bla*
	Fragment *pt_PA* was amplified from pHS_PA using primers RBSpl_tox_PA_F and pl_tx_PA_R (Table 2.3), adding a T7 ribosomal binding site using Taq polymerase. The amplicon was inserted into the pGEM T-Easy vector and sequenced to confirm construct orientation downstream of the T7 promoter. Part of the pGEM_PT series.
pGEM_PAL	P_{Lac}-*pt_PA*, pGEM T-Easy vector. High copy number. *bla*
	Fragment *pt_PA* was amplified from pHS_PA using primers RBSpl_tox_PA_F and pl_tx_PA_R (Table 2.3), adding a T7 ribosomal binding site using Taq polymerase. The amplicon was inserted into the pGEM T-Easy vector and sequenced to confirm construct orientation downstream of the P_{Lac} promoter. Part of the pGEM_PT series.

Table 2.3: Primer sequences

Primer	Sequence
B:BncoI_F	TGGAGAAAATACCATGGAACAAAGCACTAT
B:BpmeI_R	AGCTTGGGGTTTAAACTTCCATATTGTTCATCTCC
araCBAD_BB F	AGCGGCAAGCTTAATGTGC
araCBAD_BB R	TCACTTCTGAGTTCGGCATGGGG
blsc_MCS_F	GTTGTAAAACGAGGGCCCGT
blsc_MCS_R	ACTTTATGCTTCGGGCCCGTA
M13 R	CATGGTCATAGCTGTTTCC
M13 F	GTAAAACGACGGCCAGT
T7_F	TAATACGACTCACTATAGGG

Table 2.3: Primer Sequences continued ...

Primer	Sequence
ptac:bs_F	GTTTGAGACGGGCGACAG ATC C
ptac:bs_R	GAA CTG ATC GAC ACT GCT CGA GCC GC
BbpmeI_R	TGGGGTTTAAACTTCCA
ptac:bn_F	GAATTCGAGCTCGAGCT
bngene_R	CAGATTGGACCGGACTC
RBSpl_tox_PA_F	CGGCCAGTGAGCTCTTAACTTTAAGAAGGAGATATACATA CCCATGGCAACTTTGAATCCTACCAATGC
pl_tx_PA_R	TCACACAGGGAGCTCGCCGCCCTTGCTGTTTC
Bn_ncoI_F	CTGTGACAACCATGGCACAGGTTATC
Bs_ncoI_R	GGGTTTGTGTTTCCATGGTGTTCATCTC
GST F	CTTTGCCGTTAACCCTAAGGG
GST R	GCTGCAATGTGCTCTAACCC
PT_E.coli_F	CAAGGCGTTCATTGCTTCTT
PT_E.coli_R	TGACAACACAACATATTATCGAACC
PT_P.aeur_R	CTACCAATGCAATCGCAACC
PT_P.aeur_F	CTACGCGGTGAGTGACTG
PT_B.pseud_R	TTCATGACCATGCCTCGTTAG
PT_B.pseud_F	CTCGGGAAGTTATCCATCACAC
PT_A.baum_R	TCATTAGCTGCACGCTTGG
PT_A.baum_F	ACATATTATCGAACCAGGGCAAG
PT_X.camp_R	GGCTTACGCTGGGTGAA
PT_X.camp_F	AACTCAGGCAGTGCATCAT

2.3 Bacterial Growth Assays

Barnase and barstar expression

Two different procedures were used to measure the effect of barnase and barstar expression on strains CSH100 and LMG194. Growth was initially monitored by OD_{600} measurements utilizing a Fluorostar Optima plate reader (BMG Labtech). Saturated cultures derived from an overnight (o/n) culture were diluted 10^2-fold in fresh LB

and grown to ~0.100 OD_{600} in 10 ml McCartney flasks, as before. The cultures were then aliquoted (1 ml) into 24 well tissue culture plates (Jet Biofil), treated with or without inducer, either 1.0 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) or 0.00002%-0.2% arabinose (ara). Plates were placed in the plate reader, at 37°C at 200 rpm, with measurements taken every 5 mins (as the average of 20 readings), for 4 hrs. Each condition was replicated four times per trial, and at least three trials were performed. To determine the number of colony forming units (CFUs) in culture during barnase expression, cultures were grown in the 37°C water bath and sampled manually. As before, o/n cultures were diluted 10^2-fold in fresh LB and grown to ~0.100 OD_{600} and treated with or without inducer. Cells were sampled at 30 min intervals and OD_{600} measurements were taken on an Eppendorf BioPhotometer spectrophotometer. To enumerate colony forming units (CFUs) per ml, aliquots were taken at 1 hr intervals, serially diluted and plated on LB plates supplemented with ABs. Plates were incubated o/n before counting CFUs. Each condition was replicated twice per trial, and at least three trials were performed.

Barnase re-induction

The ability of barnase expression to be turned off and on in strain CSH100, and the effect on viability was measured by enumerating CFU counts per ml. Saturated cultures were diluted 10^2-fold in fresh LB and grown to ~0.100 OD_{600} in 10 ml McCartney flasks. Cells were then induced with IPTG for 2 hrs. At 3 hrs (T3), all cultures were washed twice in phosphate buffered saline (PBS) by centrifuging the cultures for 2 min at 14 krpm to pellet the cells, resuspending in PBS, centrifuging again and decanting the supernatant. Cells were resuspended in medium +/- IPTG. A subset of cultures were again exposed to inducer 3 hrs after washing (T6). Measurements were taken 2 hrs (T8) and ~18 hrs (T24) after. Each condition was replicated twice per trial, and at least three trials were performed.

Barstar Protection

The ability of barstar expression to mitigate the toxic effect of barnase expression was tested using plasmid pBB05 and strain CSH104 (C4BB05). Cultures of C4BB05 were grown to saturation with 1.0 mM IPTG to induce barstar expression. Saturated cultures were diluted 10^2-fold into fresh LB with IPTG. Cultures were grown to a

density of ~0.100 OD_{600} before induction of barnase with 0.2% ara. Cells were grown for 1 hr at 37°C at 200 rpm in a Fluorostar Optima plate reader (BMG Labtech) before treating with 0.2% ara. Cell culture density measurements were taken at 600nm (OD_{600}) every 5 min for 6 hrs. Each condition was replicated four times per trial, and at least two trials were performed.

Toxicity of PT in liquid culture

To experimentally test whether the PT and ptaRNA1 (PT-ptaRNA1) operon constituted a type I TA system, the toxicity of putative PT proteins was tested. Toxicity was measured by effect on Rel606 (R6) and MG1655 (MG) culture density. These strains were chosen for their apparent lack of a PT gene on the chromosome. Cultures (o/n) of R6 BD_PT and MG BD_PT were diluted 10^2 in fresh LB and aliquoted (1 ml) into 24 well plates. Bacteria were cultured for 1 hr at 37°C at 200 rpm in a Fluorostar Optima plate reader (BMG Labtech) before treating with 0.2% ara. Cell culture density measurements were taken at 600nm (OD_{600}) every 5 min for 6 hrs. Each condition was replicated four times per trial, and at least three trials were performed.

Toxicity of PT on solid media

High copy number vectors containing PT were quickly lost in liquid culture. This was evidenced by the rapid replacement of plasmid-bearing bacteria by plasmid-free bacteria. A modified toxicity assay was performed on solid media. To measure differences in cell growth in response to PT induction on solid media, colonies of R6GEM_PT (an *E. coli* B strain) and BLGEM_PT (an *E. coli* B strain with an inducible T7 polymerase) were streaked onto plates of RM and M9 media supplemented with glucose (inhibitory to induction), or glycerol (permissive to induction) +/- IPTG. Plates were incubated 24 hrs at 30℃. Each condition was tested in at least three independent trials.

2.4 Protein and RNA Analysis

2.4.1 Transcription analysis

RNA Extraction and DNA removal

To confirm function of the promoter systems, both synthetic (pBD_PT) and native (pHS_PT), bacteria containing these constructs were tested for PT transcripts. RNA was extracted from strain Rel606 containing plasmids series pBD_PT and pHS_PT. Cultures (o/n) of R6BD_PT and R6HS_PT were diluted 10^2 and grown to 1.00 OD_{600} before sampling. R6 pBD_PT strains were induced with 0.2% ara at 0.300 OD_{600}. Immediately upon sampling, the culture (0.5 ml) was incubated with 1 ml RNA Later (Qiagen). RNA was extracted with the Qiagen RNeasy Mini Kit. Cells were lysed following Protocol 2 for samples with 5x 10^8 cells, using 15 mg/ml of lysozyme. Total RNA was purified using Protocol 7. Residual DNA was digested on column using the Qiagen RNAse Free DNAse Set (Qiagen Protocol, Appendix B), for 30 min at 30°C. RNA was quantified with a ND-1000 Spectrophotometer nanodrop (Biolab) and 10 μg RNA was treated with an additional 5 μl DNAse (Qiagen stock, with 10 μl Buffer RDD and water to 100 μl) for 2 hrs at 30°C. The RNA was further purified using a second column to remove buffer and DNase using the RNA Cleanup Protocol.

RT-PCR

RT-PCR was performed on isolated RNA to test for the presence of PT transcripts. The procedure was carried out with the Invitrogen SuperScript One-Step RT-PCR kit, which uses a coenzyme mix for both cDNA synthesis and downstream PCR reactions. Glutathionine-S transferase primers (GST F and R) to confirm sufficient cDNA was used as a positive control. DNA contamination of the RNA samples was tested, using KAPATaq readymix (Kapa Biosystems). Primers used were PT_*E. coli* _F and R, PT_P.aeur_F and R, PT_B.pseud_F and R, PT_A.baum_F and R, and PT_X.camp_F and R. All PCR reactions were run with a 30 sec extension time, a 55℃ annealing temperature, for 25 cycles.

2.4.2 Translation analysis

Protein extraction

Samples of CSBB01 were collected to confirm the presence of intra- and extra-cellular barnase after induction. Cultures (o/n) were diluted 10^2-fold in fresh LB then grown to an OD_{600} of ~0.500 before induction with 1.0 mM IPTG. Samples (5 ml) were centrifuged to separate cell pellet from supernatant. Cell pellets were suspended into 750 μl protein lysis buffer (Appendix A) and processed for 1 min at 21 amp on a UP200S Ultrasonic Processor sonicator (Hielscher). Protein was quantified via nandrop (BioLab; ND-1000) at 280 nm. Supernatant (4 ml) was filter sterilized (0.22 mm) and proteins were precipitated by incubating supernatant with equal volumes of 20% TCA on ice o/n, then centrifuging at 11 krpm at 4°C for 20 min. Protein pellets were washed with cold acetone.

SDS PAGE

Sodium dodecyl sulfphate (SDS) polyacrylamide gel electrophoresis (PAGE) was used to visualize protein samples from barnase-expressing cultures of bacteria. Samples were suspended in Laemmli solution (Appendix A) and run on 4-20% Tris-Glycine gels in Tris-Glycine SDS Buffer (Appendix A). Gels were stained in Coomassie Blue and destained with 10% acetic acid and 5% methanol. Samples were normalized by running equal volumes on a gel and calibrating protein content using the computer program GeneTools (Syngene).

2.5 Zones of Inhibition

Zones of inhibition (ZOI) from secretion of barnase

A ZOI assay was used to determine if secreted barnase affected the growth of neighbouring bacteria, following modified methods of Ramos *et al.*. (2006). Saturated cultures (10 μl) of LMBADLacZ and LMBB02 were spotted onto plates and incubated o/n at 37°C to allow secretion of barnase into the medium. The following day, the spots were overlaid with test strain LMG194 by spraying culture from a sterilized spray bottle. This method was repeated in LB, M9, and RM media with glucose and glycerol, and with or without different concentrations of ara to induce barnase.

Overlain cultures were used at log and saturated phase to vary the concentration of bacteria placed over barnase-expressing colonies, and potentially exposing differences in growth-phase dependent barnase susceptibility. Cultures were gown on plates at either 30 or 37℃, from 24 to 48 hrs and inspected for halos around the LMBB02 spots.

ZOI from secretion of barnase with barstar expression

The ZOI assay was repeated with C4BB05, which allows induction of both barnase and barstar, reducing toxicity from excess barnase. Colonies of C4BB05 and C4BS01 were picked and grown in liquid LB + AMP and IPTG to induce barstar expression, at 37℃ for 8 hrs. Aliquots (10 μl) were spotted on LB, RM and M9 plates with glycerol, IPTG, and +/- ara to induce barnase secretion into the media. Plates were incubated at 37℃ or 30℃. On Day 2 (LB and RM media) or Day 3 (M9 media), a day culture (1:100 dilution) was made from a saturated culture of Rel606 or CSH100 and grown to mid log phase. Sterilized paper discs were dipped in the broth and gently applied to the plates of C4BB05 and C4BS01 to overlay the culture. Plates were incubated for a further 24-48 hrs. Images were taken with G:Box gel imager (Syngene).

Spotting lawns with lysed and unlysed broth culture

The ability of barnase in the supernatant or within cells to cause ZOIs was tested. Strains of LMBADLacz and LMBB02 were grown to 0.100 OD_{600}, treated +/- 0.2% ara and grown for an additional 2 hrs. Supernatant was collected, filtered, and TCA precipitated (see protein extraction above). Aliquots of whole culture were lysed by freezing 5x in liquid nitrogen for 30 sec.Strain LMG194 was cultured to 0.100 OD_{600} and spread onto M9 and RM plates with glucose or glycerol +/- ara to form a lawn. The lawn was spotted (10 μl) with the unlysed, lysed, and supernatant prepared from the LMBADLacZ and LMBB02 cultures. Plates were incubated at 30℃ for 2 days.

2.6 PSK tests

Temperature-dependent PSK assay

The ability of two PT-ptaRNA1 operons and barstar and barnase to confer PSK activity upon a plasmid was tested using the temperature-sensitive (for replication) plasmid pHSG415. This experiment used the pHSG415 derivatives described as pHS_EC, pHS_PA and pHS_BB (Table 2.3). The temperature-sensitive plasmids are unable to replicate at 42℃. Saturated cultures of bacteria were diluted 10^2-fold using LB+AB to a final volume of 2 ml and grown at 30℃ until reaching ~0.200 OD_{600}. Culture (1 ml) was concentrated by centrifugation at 13.5 g for 2 min, producing a pellet that was resuspended in PBS, centrifuged, and resuspended again (washed). Washed cells were added to fresh LB-AB (10^2 -fold dilution) and grown at 42℃ for 4 hrs. Cells were diluted 10^2 -fold periodically to maintain the culture in log phase (below 1.00 OD_{600}). The cultures were sampled every 2 hrs and plated on LB, + and -AB. Plates were incubated at 30℃, and CFUs enumerated the following day. Results were taken as the log ratio of plasmid-containing cells to all cells, such that at 0 indicates all cells had plasmids, and 1 indicates only 10% had a plasmid. Cultures with +PSK plasmids should not accumulate plasmid-free cells.

Transformation-based PSK tests

PSK-like activity of barnase-bearing plasmids was tested using a transformation-based assay (Naito *et al.* 1995). *E. coli* strain Rel606 containing the test plasmids (pLK_MCS, pBB03, pBS03, pBRT) were made chemically competent with $CaCl_2$ (method above) and frozen in 100 μl aliquots at -80℃. Aliquots were thawed on ice and 80 μl of thawed bacteria were transformed with either 100 ng pLK_CM (compatible plasmid) or 75 ng pACYC184 (incompatible plasmid- normalized by plasmid size). Transformations were carried out as above, and samples of the transformants were plated on LB+KAN (resident plasmid), LB+CM (incoming plasmid) and LB+KAN+CM (both). Plates were incubated at 37℃ and CFUs enumerated the next day.

Long term culturing PSK tests

PSK-like activity of barnase-bearing plasmids was tested by stability in monoculture. *E. coli* strain CSH100 was transformed with one or two test plasmids (pLK_MCS, pBB03, pBS03, pBRT, pBRA). Transformants were grown to saturation in the appropriate AB. Cultures were then washed once and resuspended in LB-AB. Washed cultures were initially diluted 10^3-fold in fresh media in 12 well tissue culture plates (Jet Biofil) and grown at 37℃, aerated on a rotating platform at 200 rpm. This was repeated every 24 hrs, or ~10 generations. At stated intervals, samples were serially diluted and plated on LB-AB, and LB+AB to determine percentage of culture contained no, one or both plasmids. Plates were incubated at 37℃ and CFUs were enumerated the following day.

2.7 Computational Analyses

2.7.1 Type I TA toxin family screen

Genomic and toxin family databases

Databases of translated sequence data and known toxin loci were assembled as a baseline for which to find new type I toxin families. Sequence data from phage, plasmid, and bacterial chromosomes were downloaded from the EMBL nucleotide archive (`http://www.ebi.ac.uk/ena/`). Entire genome sequences were translated in six frames, starting at nucleotide 1, 2, and 3 and in reverse, -1, -2, -3. This created a catalogue of all possible amino acid sequences derived from the genomes, which was used for subsequent searches for TA loci. Nine families of type I toxins (Hok, FlmA, Fst, Ibs, Ldr, TxpA, TisB, ShoB, and Plasmid_toxin) were chosen for analysis. Profile hidden Markov Models (HMMs) were made for each family and used to assemble a baseline compilation of loci from known families from the translated sequences (Paul Gardner and Nicole Wheeler, personal communication).

Structure based similarity scans

The integration of known type I toxins into membranes was investigated as a way of categorizing families of toxins. Amino acid sequences for the nine toxin families were aligned and filtered to 90% identity by Dr. Paul Gardner. His filter removed

redundant sequences; those more than 90% identical to sequences already in the set. Remaining non-redundant sequences from each family (Appendix B) were run through the trans-membrane hidden Markov model web server TMHMM Server v. 2.0 (`http://www.cbs.dtu.dk/services/TMHMM/`, Krogh 2001), which predicts the orientation of the protein in a membrane. Regions of the protein are predicted to be cytosolic, within the membrane, or outside the membrane. These groupings were to be used to derive HMMs aggregated from the individual family HMMs. Aggregated HMMs were searched against the database, and compared to see if novel loci were detected, i.e. were not in the catalogue of toxins found with the individual family HMM searches.

Sequence based similarity scans

Individual family HMMs were analysed and grouped by sequence similarity. To derive a tree, the family HMMs were compared with a program called Profile Comparer (PRC; (Madera 2008)), which gives a co-emission score from HMMs. These scores were entered into a distance matrix, and used to build a tree with the program PHYlogeny Inference Package (PHYLIP) 2.0 (Elsenstein n.d.). Six aggregated HMMs were made at the nodes of the tree by combining all individual sequences within the downstream families. Sequences were tree-weighted, which corrects for different numbers of sequences within each family. Aggregated HMMs were used to find novel loci as with the sequence based similarity scan (above).

Analysis of potential new families: Any loci detected by the aggregated HMMs not in the database of loci from known families were analysed. Toxin homologues were aligned along with flanking regions and inspected visually for obvious terminator sequences and potential RNA antitoxins. RNA co-fold (`http://www.e-rna.org/cofold/`, (Mathews *et al.* 1999) was used to identify any interacting regions between the opposing strands.

Cross-TA system analysis

Data on phylogenetic range of type I, II, and III toxin systems represented by detected loci, and their distribution between chromosomes and plasmids were compiled and compared to that derived from our HMMs using published previous scans of genomic data for TA systems. Annotation of loci was compared to data logged into Pfam, Rfam, and European Molecular Biology Laboratory (EMBL) nucleotide archive. For

these analyses, FlmA was concatenated with Hok due to high levels of sequence similarity.

2.7.2 Testing PT-PtaRNA1 as a type I TA system

Choosing candidate ORFs: Of the toxin families used to screen genomes, one, plasmid_Toxin (PT), has yet to be experimentally validated for toxin-antitoxin activity. To do this verification, five PT ORFs were chosen for testing toxicity and two full operons with antitoxin were tested for the PSK phenotype. Candidate loci were found by using the individual family HMM for PT, as well as compiling annotations from databases such as Pfam and Rfam. For closer analysis, five were chosen due to their location on a mobile element, availability of experimental validation, or likelihood to be functional in *E. coli*. Test strains MG1655 and Rel606 were screened for genomic regions of similarity to PT. ORFs were identified through two ORF predictors, ORF finder and EasyGene (NCBI, `http://www.ncbi.nlm.nih.gov/gorf/gorf.html` and `http://www.cbs.dtu.dk/services/EasyGene/`, respectively). Promoters were identified using BPROM, bacterial prediction software (`http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb`) (Solovyev and Salamov 2011).

2.7.3 Distribution of barstar and barnase on plasmids and viruses

A literature search was performed to find RNases and inhibitors that interact with barstar and barnase. The presence of the genes on plasmids and viruses was determined using blastp on the NCBI genomic database of all bacterial sequences. Searches were performed using the *B. amyloliquefaciens* barnase and barstar amino acid sequences as the comparators in their active form (omitting the export signals on native barnase), and repeated with amino acid sequences from the interacting RNAses and inhibtors (Table B.1). The resulting matches were compiled using the program Geneious, which filtered the entries by the terms ‘virus’ and ‘plasmid’. Results were validated as occurring on viruses or plasmids manually. Blastn was used to find matches to loci of linked barnase and barstar.

2.7.4 Programs and statistics

This thesis was typeset in LaTex. Figures and graphs were generated in Excel and LaTex, with variation measured as Standard Error (SE). P-values were generated using two-tailed t-tests in Excel. Though most conditions had equal variance between them (as established with an F-test, also done in Excel), all t-tests were performed assuming unequal variance to account for those exceptions and to keep consistency across tests. Unless otherwise stated, t-tests were performed on data from the final time point of an assay, comparing each condition back to the control.

Chapter 3

Distribution of type I TA systems across replicons, comparing between families and other TA types

Bioinformatic techniques have been instrumental in expanding the number of known families and putative homologues of TA systems. TAs have been found on a multitude of MGEs, and are abundant on microbial chromosomes, with some species of bacteria containing more than 90 copies of type II systems (Pandey and Gerdes 2005; Ramage *et al.* 2009; Leplae *et al.* 2011). Interesting patterns in distribution have emerged between types (defined by antitoxin) and between families within types (categorized by relatedness of the toxin). Thus far, type II systems appear to be more broadly dispersed than type I systems, occuring across many phyla of bacteria as well as in archae (Leplae *et al.* 2011; Makarova *et al.* 2009). In addition, type II families are more commonly localized on mobile elements and more lineage independent: that is, a given family of type II TAs appears across unrelated species of bacteria (Fozo *et al.* 2010; Leplae *et al.* 2011; Makarova *et al.* 2009; Weaver *et al.* 2009).

This differential distribution has caused some to suggests that type I systems are less horizontally mobile than type II systems (Fozo *et al.* 2010; Mruk and Kobayashi 2014). This hypothesis is investigated here. Historically, type I systems have been more difficult to detect *in silica* than type II systems, though recent computational analyses have begun to change this (Fozo *et al.* 2010; Findeiß *et al.* 2010; Sayed *et al.* 2012; Kawano 2012; Fozo *et al.* 2008). And as our ability to detect type I TAs advances, differences in distribution in comparison to type II systems may become less apparent.

Computational methods are used to explore the phylogenetic range of known families of type I TA systems and find new families, which may display different distributions than known families. These results are compared to findings in the existing literature on TA families. Aspects of TA system biology, and limitations of current methods are discussed which may account for why type I systems consistently differ in their phylogenetic range from other known TA systems. Most of the raw data in this chapter was generated through close collaboration with Dr. Paul Gardner.

3.1 Results and Discussion

Most type I toxins are small and membrane associated, (Kawano 2012; Fozo *et al.* 2008; Wang *et al.* 2012; Unoson and Wagner 2008; Mok *et al.* 2010). SymE, a nuclease that targets mRNA, is an exception (Kawano *et al.* 2007). Nine families of type I TA systems where chosen for analysis, all with known or predicted membrane-associated toxins: Hok, FlmA, LdrD, TxpA, Ibs, TisB, ShoB, Fst, and plasmid_Toxin (PT). TisB, Ibs, and Hok are believed to insert into the inner membrane, causing loss of membrane potential and the ghost cell phenotype (Unoson and Wagner 2008; Gerdes *et al.* 1986; Gerdes *et al.* 1986; Fozo *et al.* 2008; Mok *et al.* 2010). The Ldr toxin causes nucleoid condensation upon overexpression (Kawano *et al.* 2002) and the Fst toxin causes chromosomal mis-segregation and interferes with cell division at low levels as well as disrupting cellular membranes at higher concentrations (Patel and Weaver 2006). These families were to analyzed to identify patterns between families that would facilitate finding new families of type I TAs, and their distributions were used as a source of comparison to other TA types.

A database derived from six frame translations of bacterial chromosomes and plasmids was created. The database was scanned with hidden Markov models (HMMs) for each toxin family based on amino acid sequence (Figures C.1 to C.9). All regions that were considered 'matches' in the database (loci) were assembled into a baseline compilation of putative homologues from known families. We then attempted to group families by different methods to generate new HMMs, aggregated from the sequences of multiple type I families. This was intended to broaden the signal, decreasing the specificity of the HMM and increasing the ability to detect previously unknown families. Loci found with the aggregated HMMs were compared to the compilation of loci from known families.

Table 3.1: Scoring type I TA system toxins by membrane topology using TMHMM

	Total[1]	Predicted transmembrane[2]	In-Out[3]	In-Out[4]
Hok	44	39	22	17
FlmA	35	33	19	14
TxpA	1	1	0	1
ShoB	1	1	1	0
Ibs	20	1	0	1
Plasmid_toxin	11	0	0	0
Ldr	9	9	1	8
Fst	15	3	3	0
TisB	8	4	4	0

1 Number of non-redundant sequences from each family analyzed with TMHMM

2 Number of sequences that scored as transmembrane proteins

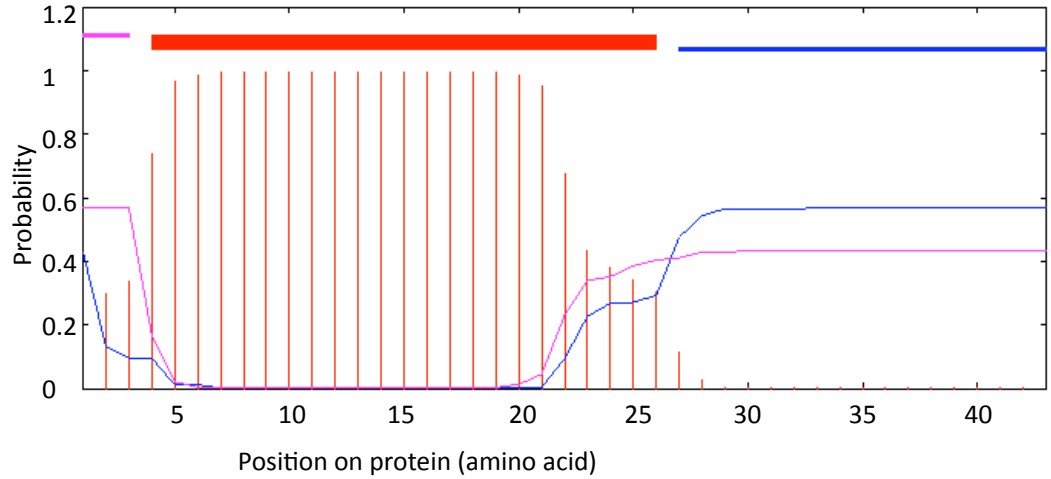
3 N terminus is inside the membrane, C terminus is out

4 N terminus is outside the membrane, C terminus is in

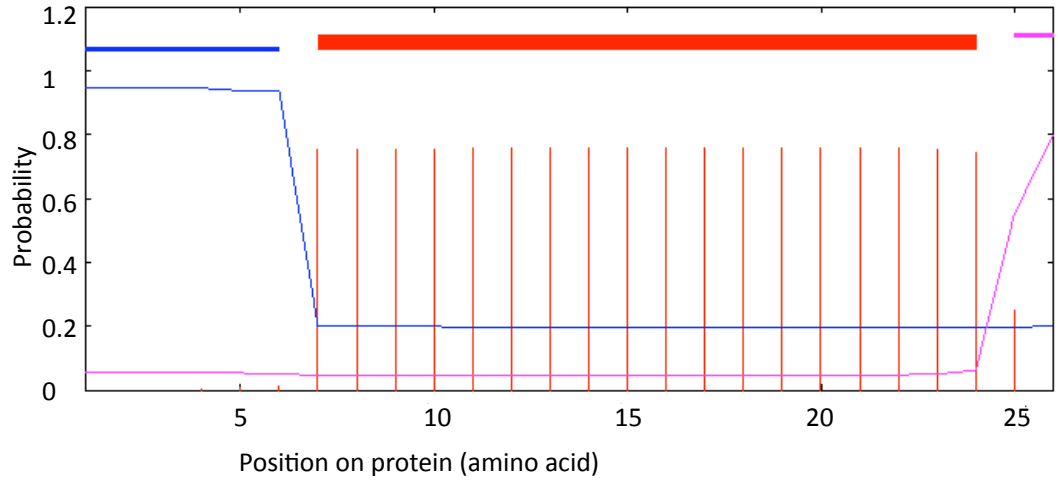
3.1.1 Structure-based searches for new TA families

First, families of toxins were categorized by the predicted structure of the putative toxins within the membranes. To this end, non-redundant amino acid sequences for nine families of type I toxins were input into the trans-membrane HMM web server, TMHMM server 2.0. The output of the server is a graph showing the probability that a given region of the protein is inside, within, or outside of the membrane (Figure 3.1). Output graphs for all TMHMM inquiries available in the e-supplement.

The TMHMM for each input sequence was scored based on the location of the N-terminus of the protein. When the N terminus was predicted to be on the cytoplasmic side of the membrane, the protein was classified as in-out, and when the N terminus was predicted to be external, the protein was classified as out-in (Table 3.1). The results from TMHMM were not consistent within families, both in prediction of transmembrane domains and whether the proteins were classified as in-out or out-in.



(a)



(b)

Figure 3.1: Membrane orientation of representative type I toxins. From the TMHMM program output, membrane orientation is given as the probability (y-axis) that a given region (x-axis, residue number) of the protein is cytosolic (pink), embedded within the membrane (red) or external (blue). A) Hok from *Proteus mirabilis* strain HI4320 (AM942759) predicting an out-in orientation. B) ShoB from *E. coli* O55:H7 strain CB9615 (CP001846) predicting an in-out orientation.

A large number of the peptides were not predicted to be transmembrane proteins by TMHMM (Table 3.1), particularly within the Ibs and PT families. The TMHMM program is known to have limitations in predicting membrane proteins. It is designed to detect transmembrane domains that are alpha helices, but does not predict the beta barrels of porins, or beta helices of gramicidin A, for example (Krogh *et al.* 2001). Though the software has been shown to be robust in comparison to other prediction software (Kall and Sonnhammer 2002), small and membrane associated proteins may be particular challenging. Some families do not have resolved structures, and how they integrate and disrupt membranes is not entirely known, even for well-studied proteins such as Hok (Gerdes *et al.* 1986; Pedersen and Gerdes 1999; Pecota *et al.* 2003).

Of those families that had a significant proportion of members that were transmembrane, not all were consistently in-out or out-in (Table 3.1). In the case of the Ldr family, all scored as an out-in sequence except for one. This one had only a weak probability (between 0.5 and 0.6) of being in the opposite orientation, having only a slightly greater than 50% chance of having an internal N-terminus. It may have been possible to group this family as out-in overall. On the other hand, there were toxins in the FlmA family that scored strongly for both orientations, making it difficult to get a definitive grouping.

The sequences we used were mainly from sequenced chromosomes (see below). As many chromosomal loci have inactivated toxins and are in a state of incremental decay (Mine *et al.* 2009), they may no longer code for residues crucial to the peptide functioning within a membrane. This, as well as differences in the selection history of TA systems when they evolve on chromosomes or plasmids, has the potential to make the sequence signal within a family ambiguous. Membrane topology did not prove to be a parameter by which we could group type I toxin families, as results were inconsistent even within families. As such, aggregated HMMs were not derived from this data.

3.1.2 Sequence-based searches for new TA families

The nine toxin families were then grouped into a tree using similarities between profile HMMs (Figure 3.3). To do this, the family HMM profiles were compared with the program Profile Comparer (PRC) (Madera 2008). This program derives

co-emission scores, the probability that similarities between the HMMs could arise by chance. These scores were entered into a distance matrix, and used to build a tree with the program PHYLogeny Inference Package (PHYLIP) 2.0 (Felsenstein 1981). New HMMs were made at each node of the tree by combining all individual sequences within the downstream families (Figures 3.3 to C.14).The database of six frame translated genomes was scanned with the six aggregated HMMs, and generated loci were compared to the database of known family loci to see if any novel systems were detected.

One family was detected by the Fst-Ldr aggregated HMM that was not one of the nine families used to derive the HMMs. It occurred primarily on plasmids and chromosomes of different *Staphylococcus aureus* strains (two to five copies), although loci were also detected in *Staphylococcus warneri*, *Staphylococcus saprophyticus*, *Staphylococcus haemolyticus*, *Staphylococcus epidermidis*, *Staphylococcus simulans*, *Staphylococcus lugdunensis*, and *Staphylococcus pseudintermedius*. Two loci were also found on *Listeria monocytogenes* and one each on *Veillonella parvula* and *Macrococcus caseolyticus*.

Non-redundant loci of the putative Fst-Ldr family were aligned (Figure 3.2) using RNA Code (Washietl *et al.* 2011). This and the flanking region were visually inspected. A terminator was identified on the reverse strand, $3'$ of the toxin peptide, suggesting a non-overlapping antitoxin as seen in type I TA families Tis-IstR, ShoB-OhsC and Zor-Orz. We used RNA co-fold to identify a large region interacting between the two strands (Figure 3.4).

In the process of analysing this system, we discovered that it had been recently described as a probable type I system (Sayed *et al.* 2011; Sayed *et al.* 2012). Labelled as SprA1, the RNA antitoxin was originally identified by bioinformatic searches and transcriptomic analysis (Sayed *et al.* 2011). The toxin SprA1 and antitoxin Spr1AS were shown to be constitutively co-transcribed using Northern blot (Sayed *et al.* 2011), with half lives of greater than three hours and ten minutes, respectively (Sayed *et al.* 2012). The SprA1 protein, PepA1, localized to cellular membranes, having a toxic effect on *S. aureus* and human cells (Sayed *et al.* 2011; Sayed *et al.* 2012).

Although our aggregated HMMs were unable to detect fully novel type I systems, they did enable us to detect non-input systems. This method of aggregating HMMs could be repeated with more known families. The recently discovered family SprA1 is found on plasmids, increasing the number of type I families that are mobile. As

Figure 3.2: Alignment of putative type I toxin family loci. Non-redundant sequences of genes from the putative Fst-Ldr family were aligned using RNAcode. Consensus amino acids are displayed in brown at the top, with the relative level of conservation in grey at the bottom. Codons with one to three synonymous mutations are indicated with increasing saturation of green and codons with one to three non-synonymous mutations are indicated with increasing saturation of red.

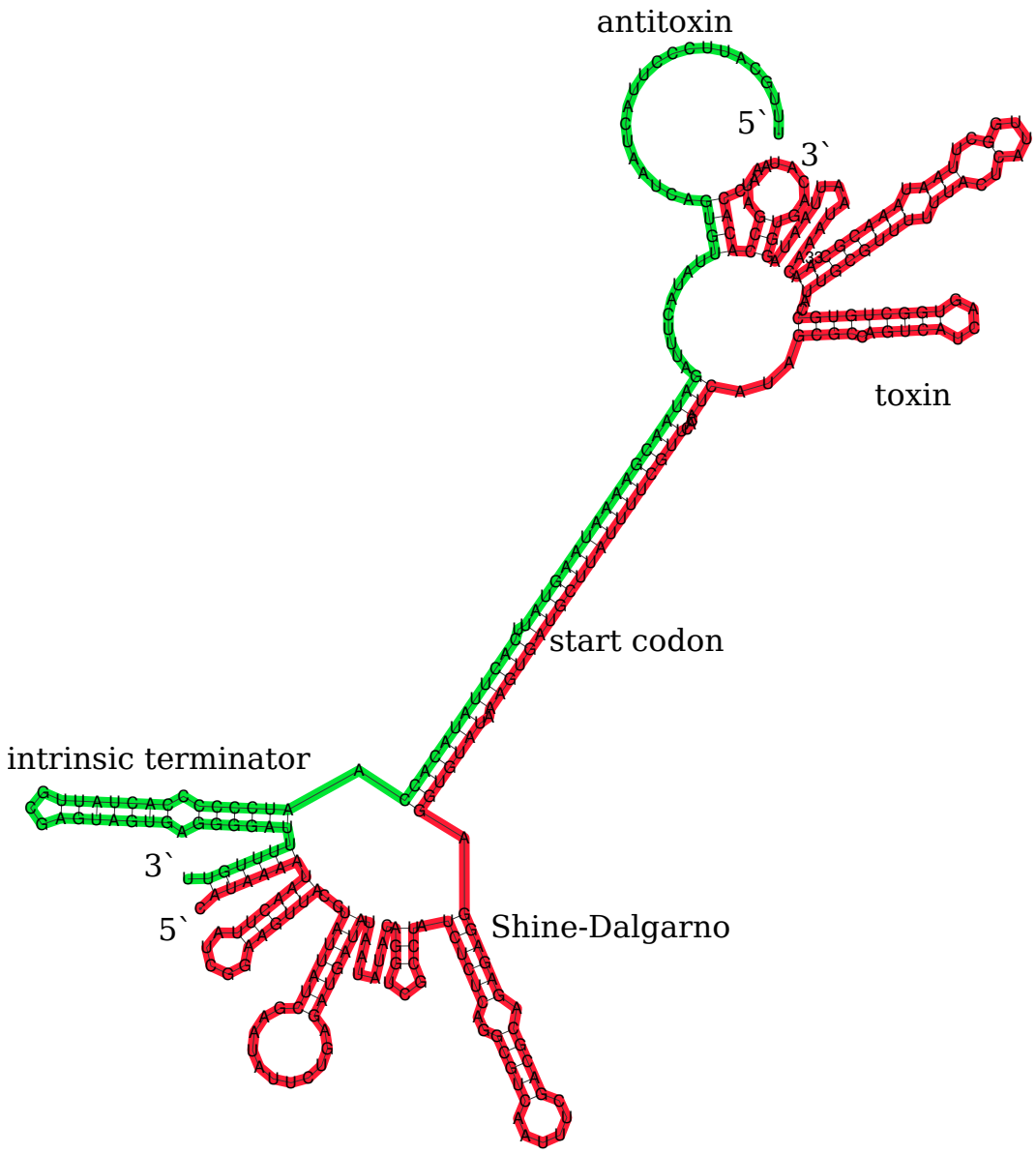


Figure 3.3: Predicted base pairing between putative type I toxin mRNA and downstream antisense strand. Interactions between the toxin (red) and antitoxin (green) were predicted using RNA Code, using the predicted loci from *S. epidermidis* ATCC 12228 (GenBank: AE015929.1).

the catalog of known type I families expands, more such mobile families may become evident.

3.1.3 Distribution of loci from known families

The process of creating a database of 'known' loci from described families revealed many unannotated putative toxins. The total number of homologous sequences was anywhere from three to six times the number reported in previous screens (Fozo *et al.* 2010) for ShoB, TisB, Ibs, and Ldr/Fst systems, and almost ten times for Hok-type systems. The opposite was true for TxpA, which was found only twice in our screen but 48 times by Fozo *et al.* (2010).

Despite the large number of new loci found in the toxin-family based searches, their phylogenetic distribution was narrow (Table 3.2). All of the Gram-negative type I systems were limited to the Proteobacteria, and Ibs, Ldr, ShoB, and TisB were all limited to just the Enterobacteriales. Both of the Gram-positive type I systems were limited to the classes Bacillales and/or Lactobacillales in Firmicutes. Many type II systems are found across phyla, some in as many as 11 or 12 eubacterial phyla, as well as in archaea (Leplae *et al.* 2011) (Table 3.2). All three type III systems were found in at least four phyla, both Gram-negative and Gram-positive (Blower *et al.* 2012). TA systems are grouped by toxin, as toxins pair with multiple families of antitoxins and vice versa (Guglielmini and Van Melderen 2012; Leplae *et al.* 2011; Chopra *et al.* 2013).

The distribution of type I modules was also investigated on plasmids and phages (Figure 3.5). Again, our findings were consistent with previous reports, with Hok, Fst, and PT on plasmids and the remaining five groups not (Fozo *et al.* 2008; Findeiß *et al.* 2010; Kwong *et al.* 2010). Some of the chromosomal matches may be on integrated elements, as was the case for the PT locus in *Burkholderia pseudomallei* (strain K96243). Most type II families can be found on mobile elements such as phage and plasmids (Leplae *et al.* 2011) (suplementary Table 2), as well as two of the three known type III families (Blower *et al.* 2012) (Table 2). This makes the distribution of type I systems across both species and replicon type comparatively narrow.

Some of the less broadly distributed type I systems investigated, such as TisB, are regulated by the SOS response and have a role in persister cell formation (Brantl 2012; Dörr *et al.* 2009). Ones without a known function may provide protection against

Table 3.2: Phylogenetic distribution of type I toxin families.[1]

Toxin	Discovery	Phyla[2]	Orders[2]
Plasmid_toxin	Bioinformatic (Findeiß *et al.* 2010)	1	7
Fst	Plasmid stability (Weaver *et al.* 1996)	1	6
Hok	Plasmid stability (Gerdes *et al.* 1986)	1	4
Ibs	Repeats in sequence data (Fozo *et al.* 2008)	1	1
Ldr	Repeats in sequence data (Kawano *et al.* 2002)	1	1
TxpA	sRNA searches (Silvaggi *et al.* 2005)	1	1
TisB	Bioinformatic based sRNA searches (Vogel *et al.* 2004)	1	1
ShoB	Cloning based sRNA searches (Kawano *et al.* 2005; Fozo *et al.* 2008)	1	1

1 Families found on a plasmids or phage are shaded.

2 Number of phyla the loci are found on, out of 12, and number of orders within that phylum.

Table 3.3: Phylogenetic distribution of type II toxin families. [1]

Toxin	Discovery	Phyla[2]
Doc	Plasmid stability (Lehnherr *et al.* 1993)	12
CcdB/ MazF	Plasmid stability (Ogura and Hiraga 1983)	11
ParE/ RelE	Plasmid stability (Johnson *et al.* 1996; Gotfredsen and Gerdes 1998)	11
VapC	Genomics/function (Zhang *et al.* 2004)	11
HipA	Persistence (Moyed and Bertrand 1983)	9
GinB	Bioinformatic screen, guilty by association (Leplae *et al.* 2011)	7
GinA	Bioinformatic screen, guilty by association (Leplae *et al.* 2011)	5
YafO	Bioinformatic screen, antitoxin homology (Brown and Shaw 2003)	1
GinC	Bioinformatic screen, guilty by association (Leplae *et al.* 2011)	1
GinD	Bioinformatic screen, guilty by association (Leplae *et al.* 2011)	1

1 Families found on a plasmids or phage are shaded.

2 Number of phyla the loci are found on, out of 12.

Table 3.4: Phylogenetic distribution of type III toxin families. [1]

Toxin	Discovery	Phyla[2]
TenpN	Bioinformatic screen, protein structure (Blower *et al.* 2012)	5
ToxN	Sequence similarity to Abi proteins (Fineran *et al.* 2009)	4
CptN	Bioinformatic screen, protein structure (Blower *et al.* 2012)	4

1 Families found on a plasmids or phage are shaded.

2 Number of phyla the loci are found on, out of 12.

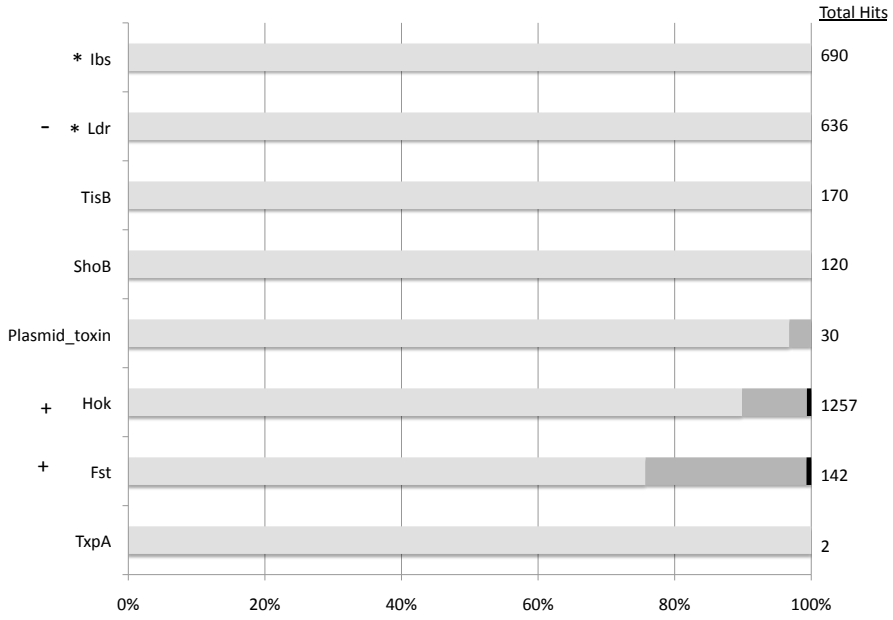


Figure 3.4: Replicon distribution of type I TA system loci. Genomic data was scanned for loci from nine families of type I toxins. Distribution of the loci across replicon type is shown here, on chromosomes (light grey), plasmids (dark grey) and phage (black). Families tested for PSK have a + (induce PSK) or – (don't induce PSK). Families occurring as tandem repeats marked with a $*$.

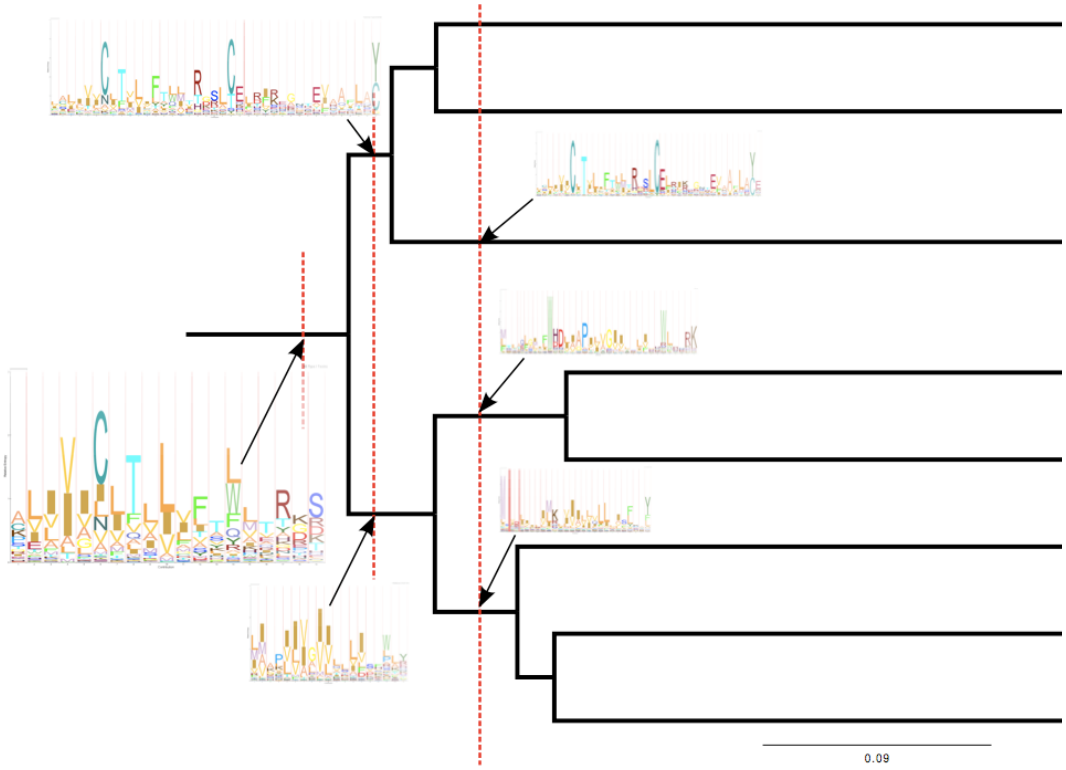


Figure 3.5: Aggregated HMMs derived from type I TA system toxins. HMMs were derived for nine families of type I TA toxins. These HMMs were grouped into a tree using PHYLIP 2.0. At each node, all downstream HMMs were combined to derive aggregated HMMs.

mobile elements such as phages (Fineran *et al.* 2009; Kawano *et al.* 2002). Any given system may be reatined by acquiring a role in cellular physiology. However, those roles are not mutally exclusive with selection and integration into new replicons. Such a role in physiology may offer an explanation of why type I systems are vertically transmitted but not why they appear to be less broadly distributed on horizontally mobile elements. It is interesting then, to ask why such differences may arise. Are they due to inherent differences in biology between types of TA systems? Could they arise from artifacts of limited datasets?

3.1.4 Phylogenetic distribution, PSK and presence on mobile elements

I am ultimately interested in the differences between types of TA systems. TA systems with similar antitoxins and sequences are also similar in their phylogenetic range and their tendency to be on mobile elements. Type I TAs are more narrowly distributed than type II and type III TA systems (Fozo *et al.* 2010; Leplae *et al.* 2011; Blower *et al.* 2012). Among type I TAs, the Hok-Sok and Fst-RNAII systems are more widely distributed than Tis-IstR and ShoB-OhsC. And, of five sequence-similar clusters within the type II MazF toxin family, some clusters plasmid-borne modules and some were strictly chromosomal (Chopra *et al.* 2013). We are looking at the interface of three traits, phylogenetic range, mobility, and PSK, attempting to determine if there is an explanation for observations of differences within families and types.

Presence on mobile elements and PSK

Many type I TA families have not been found on mobile elements (Figure 3.4). It is evident that PSK is often selected for on mobile elements, suggesting a correlation

between TAs that are highly mobile and TAs that exhibit PSK. Most TA homologues on plasmids and mobile elements tested to date are able to mediate PSK (Milunovic *et al.* 2014; Ogura and Hiraga 1983; Lehnherr *et al.* 1993; Wozniak and Waldor 2009; Rowe-Magnus *et al.* 2003; Christensen-Dalsgaard and Gerdes 2006). Mine *et al.*'s analysis of plasmid encoded *ccdF* (type II TAs) found evidence of strong selective pressures for maintaining addiction. This selection on mobile elements increases the movement of the genes between cells, affecting their spread to new loci on chromosomes as well.

There is potential for a PSK-like mechanism to maintain genes on chromosomes as well as plasmids, via competition during homologous recombination between resident TA containing regions and incoming –TA regions (Mruk and Kobayashi 2014). Addictive genes are difficult to replace by gene conversion on chromosomes because the toxin would kill the recombinant (Van Melderen and De Bast 2009; Gerdes and Wagner 2007; Tsilibaris et al. 2007; Rowe-Magnus, Guerout, et al. 2003; Mruk and Kobayashi 2014). This is true of both RM and TA systems. This creates PSK-like sele ctive pressure for their maintenance regardless of their ability to aid the host (Van Melderen and De Bast 2009; Leplae et al. 2011). This is especially likely in situations with high levels of homologous recombination (Mruk and Kobayashi 2014; Yahara et al. 2012) , and would apply to TAs on plasmids as well as chromosomes. That said, many chromosomal TA systems appear to be in a state of incremental decay (Mine et al. 2009; Pedersen and Gerdes 1999), with the toxins being inactivated first followed by regulatory regions and the antitoxin. Five of six cop ies of hok - sok on the E. coli K12 chromosome have insertions or point mutations, and none demonstrated PSK activity (Pedersen and Gerdes 1999). The study of ccd homologues found that thirty percent of chromosomally encoded toxins were inactiv ated, with evidence suggesting the chrom osomal toxins were under neutral selection (Mine et al. 2009). Of the systems that still have an active toxin, not all can mediate PSK when placed on plasmids (O'Neill, A. Chen, and Murray 1997; Kawano 2012).
Plasmid -borne type I systems tested so far have exhibited at least some features of the PSK phenotype (K. E. Weaver, Reddy, et al. 2009; Gerdes, Bech, et al. 1986). Most chromosome- specific type I systems have not been tested for their ability to function as a PSK on plasmids (Figure 5), except Ldr-RdlD, which did not exhibit PSK (Kawano, Oshima, et al. 2002 ; Kawano 2012). It is difficult to determine if

failure to observe a PSK phenotype when placed on plasmids is an inherent feature of chromosomal-only TA families, or simply that a PSK-functionality was not selected for with a particular pair of genes. This said, it may be that those type I families that are chromosome- specific are unable to induce PSK, and do not appear to be enriched on mobile elements because of the lack of selective pressure for them to be so.

Presence on mobile elements and phylogenetic distribution

The distribution of type I TAs on chromosomes tends to be more lineage dependent than type II TAs (Table 3.2). Families of all types of TAs found on mobile elements are more likely to be phylogenetically diverse (Table 3.2) than other families within the same type. This is true when comparing type I TA families, where mobile type I families are seen in more orders than non-mobile families. Yet they are still found in only one phylum, much fewer than mobile type II families (Table 3.2). This is not a complete correlation. One type III family, *cptIN*, was found to have a widespread distribution, but has not yet been seen on plasmids and phage (Blower *et al.* 2012).

There are some difficulties with inferring mobility from phylogenetic data seen here. Many apparently chromosomal loci may be on integrated mobile elements, and gaps in the plasmid sequence database could mean that the genes have simply not been found yet (Frost *et al.* 2005). Furthermore, there is a difference between transfer and transmission- a mobile element may be able to transfer into new cells but not replicate vertically. Plasmids have different host ranges (Pukall *et al.* 1996; Clark and Warren 1979) for vertical tranmission, some being more broad than others. But this does not reflect their transfer ranges (Heinemann 1991).

MGEs provide a means for genes to move from cell to cell and species to species. PSK is selectively favored on MGEs, and genes with this phenotype may be more likely to move with them to different species. It is clear that there is nothing definitional about type I TAs that prevents them from exhibiting PSK, as some families have been shown to do so, namely Fst-RNAII and Hok-Sok. But it could be that they exhibit PSK in a more narrow range of environments than other types of TA systems. If so, they would be less likely to be selected for across a wide phylogenetic range.

3.1.5 Factors that may affect expression of PSK from type I TA systems

A PSK system with a narrow range would be one where PSK was only realized in some cells. This could be due to activities of the toxin and antitoxin. The defining characteristic of TA systems categorized thus far is their antitoxin, which might naturally account for consistent differences between the types. Type I systems use a short non-coding RNA to prevent toxicity. As a constitutively expressed RNA, it may exert an off-target toxic affect if it binds to non-target mRNAs (Jackson *et al.* 2003). Since type I antitoxins are generally short, they could have a greater likelihood of binding non-target sequences in new genomes. Extensive sequence similarity may even not be necessary (Jackson *et al.* 2003; Birmingham *et al.* 2006).

Additionally, the toxin may affect the range of these systems. If the genes move horizontally, and are involved in the stable maintenance of mobile genetic elements, the toxin should target cellular functions that are highly conserved across species. This has been proposed as a reason that type II systems, which can be highly mobile, often target conserved and essential biochemical pathways such as translation (Leplae *et al.* 2011; Goeders and Van Melderen 2014). Some type I systems are toxic in non-related species when expressed at high levels, but may have more specific mechanisms of action when expressed under their native promoter or in a particular genetic background. Though generally predicted to be membrane-associated, the exact interaction many of these toxins have in the cell is unknown. Thus, the nature of the toxin or antitoxin could reduce the chances of successful vertical transmission of newly acquired type I systems. The distribution of the type I toxin SymE, which is a nuclease, would be a good point of comparison to the membrane associated type I toxins analyzed here.

Finally, how type I systems are regulated may effect PSK. TA systems require close regulation of the toxin and antitoxin. For example, type II systems are autoregulated by protein antitoxins, which interact with the toxin and act as transcriptional repressors (Mruk and Kobayashi 2014; Zhang *et al.* 2003; Kedzierska *et al.* 2007; Cataudella *et al.* 2013; Cataudella *et al.* 2012; Afif *et al.* 2001). Changes in stoichiometry between toxin and antitoxin can de-repress the operon and increase toxin transcription and subsequent expression. In contrast, type I antitoxins control toxin expression at the translational level, binding the toxin mRNA and blocking the ri-

bosomal binding site (Darfeuille *et al.* 2007; Kawano *et al.* 2007; Fozo *et al.* 2008) and/or targeting the mRNA for degradation (Fozo *et al.* 2008). However, the molecular details have not been fully determined for many systems (Mruk and Kobayashi 2014). Perturbation of the toxin and antitoxin ratio in these systems can lead to host killing or stasis. This is seen during the stress response with chromosomal type II systems. When translation rates in the cell diminish, the toxin-antitoxin ratio of chromosomal type II systems is perturbed (Christensen *et al.* 2004; Gerdes 2000), which liberates the toxin and possibly contributes to the resulting bacteriostasis (Pedersen *et al.* 2003; Christensen-Dalsgaard and Gerdes 2006; Christensen *et al.* 2003).

Regulation of type I TA systems may vary between host species. Fozo (2012) has noted that plasmid encoded type I families have antitoxins that bind the 5′ or 3′ UTR of the gene, while chromosomal antitoxins can be within the coding region (Ibs-Sib) or be divergently transcribed (Tis-IstR, ShoB-OhsC, Zor-Orz). These divergently transcribed systems do not have overlapping DNA sequences and may have less base-pairing potential; their distribution is also especially narrow.

It may be possible to test whether or not some TA systems have a greater propensity to exhibit PSK in a larger number of hosts than others. The idea that toxins that target more conserved mechanisms are more frequently mobilized could be tested using computational methods. There are also a number of chromosomal-only type I TA families that have yet to be tested for PSK.

3.1.6 Biases within the databases

There are some biases within the DNA sequence databases the affect the data. First, not all bacterial phyla are equally well represented. Both type I and type II systems are disproportionately reported in Proteobacteria and Firmicutes, the two most studied phyla of bacteria with the greatest number of sequenced genomes. Second, the dearth of plasmid and phage sequence data (Frost *et al.* 2005) can result in artificially low hits on mobile elements. Chromosomal data can only tell us so much about the mobility of a given gene, omitting genes that are unlikely to be stably inherited vertically regardless of their abundance on mobile replicons. Phylogenetic analysis can distinguish between genes that were lost or gained after species/strain divergence, but not between transient or stable genes on a recent timescale.

Bioinformatic techniques also have a bias. The use of sequence and structure information derived from a description of known TA systems to screen for new TA systems imposes an obvious bias onto the search. Many of the best-studied type II systems were discovered due to a phenotype (Table 3.2), usually plasmid stability in monoculture, as a surrogate to measure PSK functionality. On the other hand, many of the type I systems discovered recently were found on chromosomes, looking through sequence data or by searching for RNAs. It is not surprising that TA families found on plasmids initially often show up on plasmids and phage during screens. It is interesting to note that some more recent type II families discovered through bioinformatic screens of chromosomal sequences also appear to be more lineage-dependent and thus less distributed than other type II families. Type II toxin superfamilies YafO, GinC and GinD are only found on chromosomes, and in only one phylum (Leplae *et al.* 2011) (Table 3.2). There is room, then, for structure and sequence independent methods of finding new TA systems. A recent screen by Sberro *et al.* (2013) used data on difficult to clone genomic regions to find toxins.

3.2 Conclusion

Screens for TA systems to date seem to show a difference in the phylogenetic distributions of type I and type II families. There is an apparent correlation between distribution, mobility and ability to induce PSK. This raises the possibility that type I systems are less likely to exhibit PSK, reducing their persistence at new locations or on new replicons and potentially reducing their mobility. Other genetic systems are capable of exhibiting PSK, such as restriction modification (RM) systems (Naito *et al.* 1995; Nakayama and Kobayashi 1998). There are differences between the distributions of different types as well, with Type II RM systems appearing more mobile than Type I and Type III (Mruk and Kobayashi 2014), though a fair degree of domain swapping is apparent (Naderer *et al.* 2002; O'Sullivan *et al.* 2000). This could be due to similar reasons proposed for TA systems, that *PSK is an emergent property of some systems in some environments, affecting their horizontal mobility.*

On the other hand, there has been an explosion in the discovered biochemical diversity of TA systems in the last decade. In the last year alone, two new types have been proposed, with unique antitoxin activity (Masuda *et al.* 2012; Wang *et al.* 2012). A number of new families within the types have been described (Leplae *et al.* 2011; Fozo *et al.* 2008; Fozo *et al.* 2010; Findeiß *et al.* 2010), and new functions have been proposed (Brantl 2012; Pimentel *et al.* 2014). It is possible that, as our knowledge

in this area expands, we will find that certain types of TA systems are inherently more lineage-specific than others. Or, we may find more families within the types that defy current patterns, such that families within each type have either broad or narrow phylogenetic distributions. Though it was not an entirely novel type I system, our screen found the SprA1 TA system, which has been recently described and is mobile.

Chapter 4

Experimental analysis of candidate type I TA system PT-ptaRNA1

While a recent proliferation of sequence data has enabled us to both describe the distribution of known TA systems and to identify potentially novel systems, experimental validation is still necessary for confirmation of their function(s). In this chapter I will describe tests of a putative type I TA system identified using computational methods.

Findeiß *et al.* (2010) identified a small RNA from pyrosequencing data of *Xanthomonas campestris* pv. vesicatoria strain 85-10, labeled ptaRNA1 (plasmid transferred anti-sense RNA). PtaRNA1 is proposed to be part of a type I TA system. The PtaRNA1 is antisense to an ORF that, according to the program MEMSAT3, appears to code for a protein with a trans-membrane domain, dubbed plasmid_Toxin (PT) (Findeiß *et al.* 2010). The suspected ribosomal binding site for this protein is entirely overlapped by PtaRNA1, suggesting the RNA may bind this region as a means of regulation, seen with other type I TA systems (Darfeuille *et al.* 2007; Kawano *et al.* 2007; Fozo *et al.* 2008). Expression and size of the RNA was verified by northern blot. Interestingly, the distribution of the operon (PT-ptaRNA1) is lineage independent, occurring across diverse species (Findeiß *et al.* 2010). This is more common in type II systems than other type I systems (Fozo *et al.* 2010). The authors suggested that the distribution was a result of widespread HGT.

I analyzed this system in greater detail experimentally to see if it would function as a type I TA system. We examined whether the system exhibited the biochemical requirements of a TA system, particularly if the proposed toxin ORFs were in fact toxic upon expression, and if the candidate system induced PSK. This provides us

with a chance to test the validity of the bioinformatic-based prediction, based on gene organization and distribution. I also test if the distribution of this novel system is reflected in an ability to exhibit PSK and gain a selective advantage from HGT. Our previous screen for type I toxins in genomic data included PT (Chapter 3). Of the potential PT ORFs identified, we tested five candidate PT genes from different species for toxicity in *E. coli*. Two full operons of PT-RNA1 were synthesized and tested for PSK.

4.1 Results

A potential new family of type I TA systems, PT-RNA1, was discovered with computational methods by (Findeiß *et al.* 2010). Distributed across divergent species of bacteria and on mobile elements, PT-RNA1 was a strong candidate for being a horizontally mobile TA system, and exhibiting PSK. Potential loci of PT-RNA1 were identified by scanning translated genomes with a PT family HMM (Chapter 5), as well as consulting databases such as Pfam and Rfam.

Of 15 PT loci identified by the family HMM, 14 were chromosomal and one was plasmid-borne. Four were chosen for closer analysis (Table 4.1), including the originally identified locus in *X. campestris* (the only one shown to be transcribed (Findeiß *et al.* 2010)) and two associated with mobile elements, the *Burkholderia pseudomallei* locus found inside a potential integrated element, and the plasmid locus pMATVIM-7 from *Pseudomonas aeruginosa*. A fifth, the *E. coli* OK1357 locus, was not located during the family HMM screen, but was identified on the Pfam website as part of the PT family (PF12703). This locus was chosen due to the greater chance it would be expressed in the test organisms, *E. coli* wild-type strain MG1655 and *E. coli* B strain Rel606. These strains represent two major lineages of laboratory *E. coli*. The test strains are both fully sequenced, and neither contained a candidate PT-ptaRNA1 locus. This was important, as chromosomal antitoxins can negate plasmid-borne toxicity and PSK.

The full PT protein was predicted by submitting the locus and surrounding regions to both ORF Finder and EasyGene. Three PT loci, from *A. baumannii* (PT_AB), *B. pseudomallei* (PT_BP) and *X. campestris* (PT_XC) were commercially synthesized with a T7 ribosomal binding site (RBS) and cloned using the vector pBAD33

Table 4.1: Location and coordinates of plasmid_Toxin ORFs synthesized for experimental validation

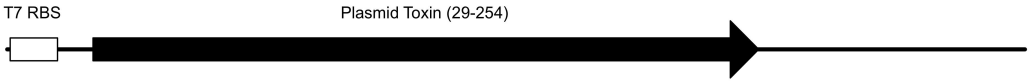
Species	Abbr	Accession Number	E-value	Bit Score	Coordinates (bp)
Acinetobacter baumannii ATCC 17978	AB	CP000521	8.8 x 10^{-39}	120.3	792285- 792501
B. pseudomallei strain K96243	BP	BX571965	9 x 10^{-36}	74	3877451- 3877235
E. coli OK1357	EC	ADUR01000110	NA	NA	594- 818
P. aeruginosa plasmid pMATVIM-7	PA	AM778842	2.9 x 10^{-37}	115.4	16679- 16457
X. campestris pv. vesicatoria strain 85-10	XC	AM039952	3.2x 10^{-39}	74	2484375- 2484591

(Figure 4.1a). The remaining two loci, from *E. coli* (PR1_EC) and *P. aeruginosa* (PR1_PA), were synthesized as operons and cloned using the vector pHSG415 (Figure 4.1b). The promoter region for ptaRNA1 was predicted for all constructs using promoter prediction software, BPROM. They were consistent with the predictions of (Findeiß *et al.* 2010). After acquiring the synthesized constructs, predicted PT genes from the two operons were cloned using the T7 RBS and inserted into pBAD33.

4.1.1 Heterologous expression of candidate type I toxins under P_{BAD} promoter on a low copy number vector

The ability of the PT constructs to express a toxic protein was first tested in pBAD33. This plasmid is a medium copy number plasmid (~20 copies/chromosome equivalent) that allows for expression under the P_{BAD} promoter (Guzman *et al.* 1995). It is commonly used for expressing toxins both because it is a lower copy number plasmid and because a strong repressor regulates the promoter. It has been used successfully to test potential TA system toxins from previous screens (Leplae *et al.* 2011; Fozo *et al.* 2010). The *E. coli* strains Rel606 and MG1655 were transformed with the constructs to test the affect of PT expression on culture growth.

(a)



(b)

Figure 4.1: Organization of synthesized plasmid_Toxin and PT-RNA1 operon sequences. Representative schematics for sequences synthesized from genomic searches. (a) *A. baumanii* PT behind the T7 ribosomal binding site (344bp), inserted 5' of the P_{BAD} promoter of pBAD33. (b) *P. aeruginosa* pMATVIM-7 plasmid_Toxin and PT-RNA1 operon (586bp) inserted into pHSG415.

The constructs were initially grown and tested in RM media where glucose maintains repression of the P_{BAD} promoter. The media was switched to LB for consistency when it became apparent that tight repression was not necessary for normal growth of the cultures (data not shown). All experiments as reported here were conducted in LB. Density of the culture was monitored by a spectrophotometric plate reader. Arabinose was added when cultures reached ~0.100 OD_{600} to induce PT expression.

The addition of arabinose did not cause a reduction in growth rate of either MG1655 or Rel606 containing PT constructs (Figures 4.2 to 4.7), suggesting that the genes do not express a product toxic to *E. coli* under these conditions. Increasing the arabinose to 2% had no affect on the growth of the PT producing strains (data not shown). A plasmid expressing a known toxin (barnase) under the P_{BAD} promoter was used to confirm functionality of the inducer under these conditions (data not shown).

In a number of cases, the addition of arabinose caused an increase in the saturation of bacterial cultures. This was most dramatic with the *A. baumanii* ORF expressed in Rel606 (Figure 4.2). This effect was not consistent across the two strain types, with PT_AB not causing an increase in growth in MG1655 upon induction. The

Table 4.2: Differences in growth between cultures induced and uninduced for PT expression in two *E. coli* strains MG1655 and Rel606, reported as p-values

Construct	MG1655	Rel606
pBD_AB	0.956	0.000 ***
pBD_BP	0.000 ***	0.042 *
pBD_XC	0.012 *	0.584
pBD_EC	0.075	0.835
pBD_PA	0.025 *	0.969
pBAD33	0.281	0.062

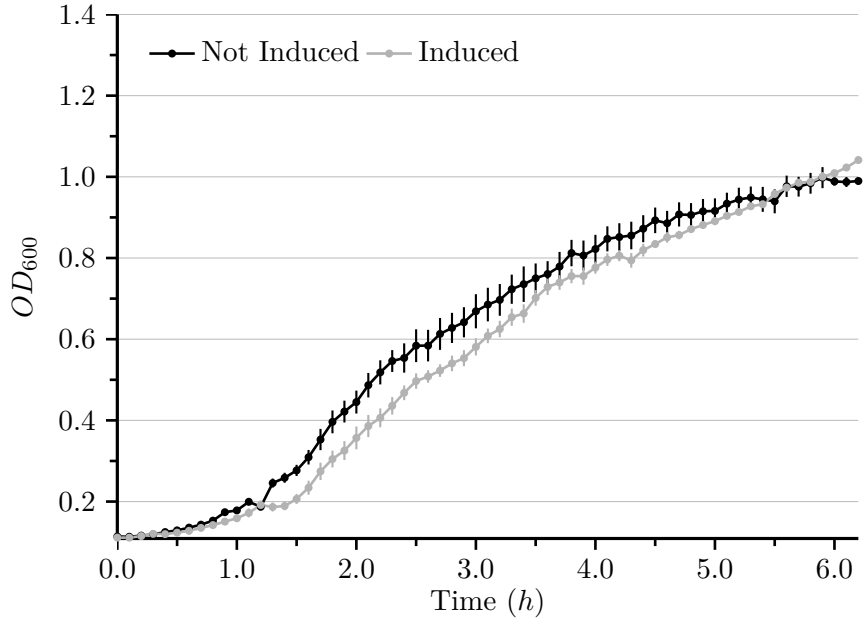
1. P-values derived from a two-tailed t-test, assuming unequal variance, on data at six hours. Results significantly different from the uninduced control are marked ($*, **, ***$) when $p \leqslant 0.05$, 0.001. 0.0001, resp.

effect of induction as a whole differed by strain and construct (Table 4.2). Overall, a growth increase was seen upon induction of more constructs in strain MG1655 than was seen in strain Rel606. With the exception of pBD_BP, induction of a given construct only significantly increased growth in one strain. As arabinose is a sugar, it could be increasing growth by increasing nutrition. This is seen in the control plasmid pBAD33 (Figure 4.7) for strain Rel606, which had the least increase in growth overall but not with the same construct in strain MG1655.

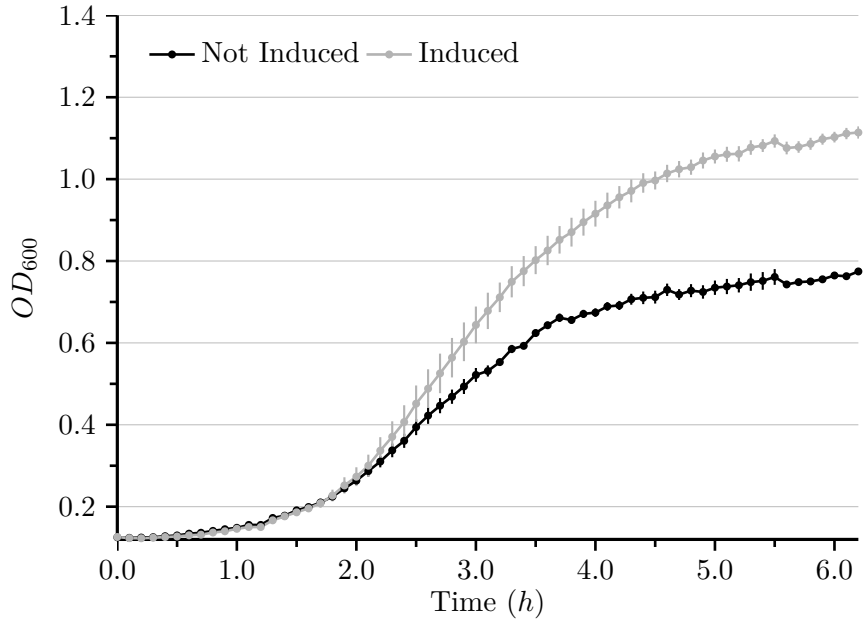
The lack of a toxic effect following induction raises the possibility that the genes were not being expressed in the system. The plasmids were sequenced again to confirm that no mutations occurred during growth and handling that would explain a lack of expression. Though the genes were synthesized under a standard promoter and RBS known to function in *E. coli*, the constructs were tested for transcription of the PT gene to provide evidence that the system was functional.

Transcription of candidate type I toxins from P_{BAD}

The small size and hydrophobic properties of the PT proteins makes it difficult to measure them using protein gel electrophoresis, so analysis was only carried out at the transcriptional level. To confirm that the promoter system was functional, the synthesized pBAD_PT constructs (pBAD_AB, BP and XC, Figure 4.8) and cloned constructs (pBAD_EC and PA, Figure 4.9) were tested for transcription. This was done only in Rel606, as both strains had similar response to PT induction. Rel606

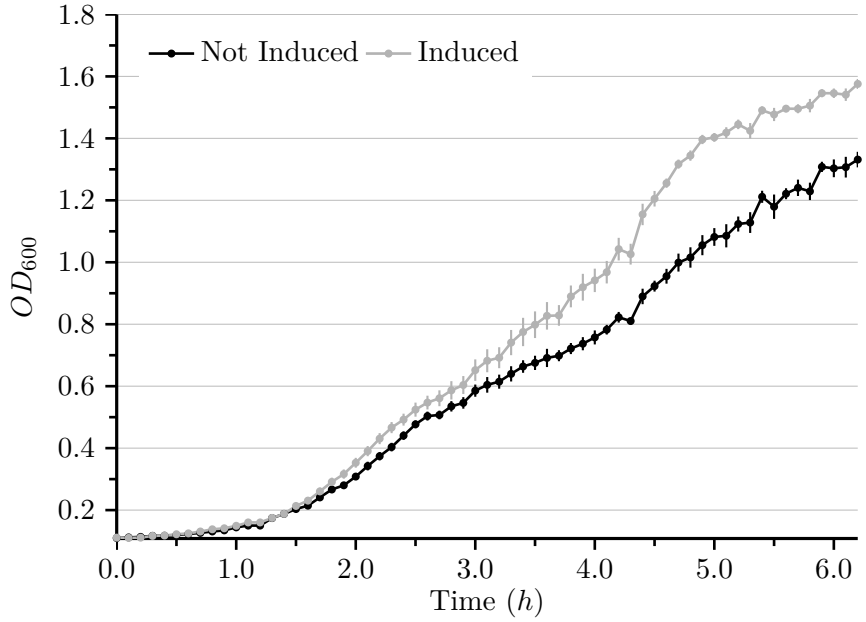


(a) MG1655

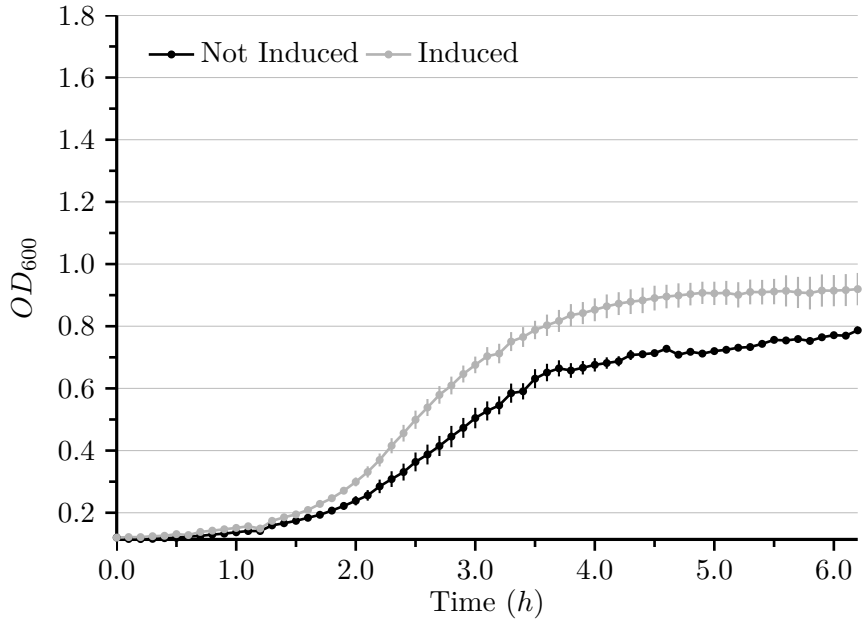


(b) Rel606

Figure 4.2: Induction of potential toxin plasmid_Toxin from *Actinobacter baumannii*, under the P_{BAD} promoter Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (black) or without (gray) 0.2% arabinose. Constructs were tested in (a) *E. coli* strain MG1655 and (b) *E. coli* strain Rel606. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.

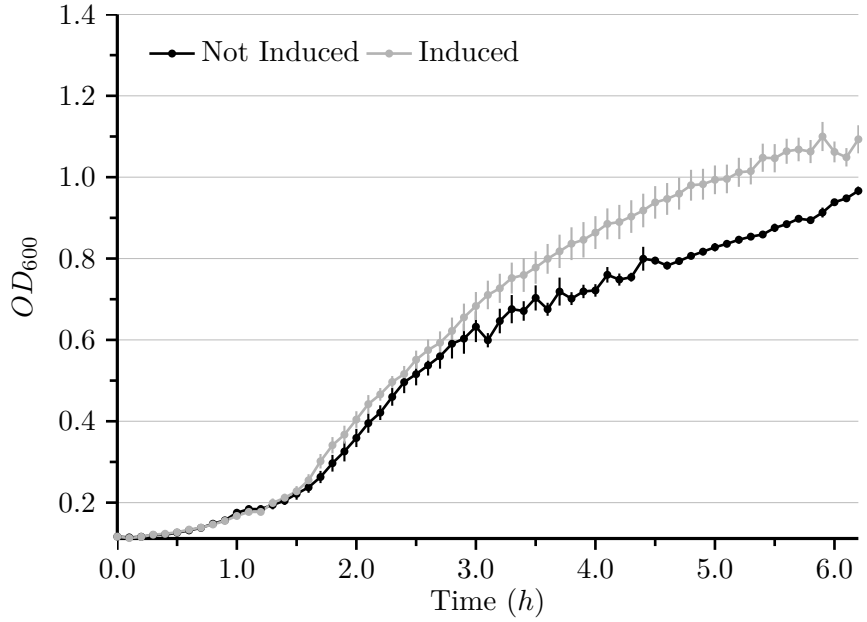


(a) MG1655

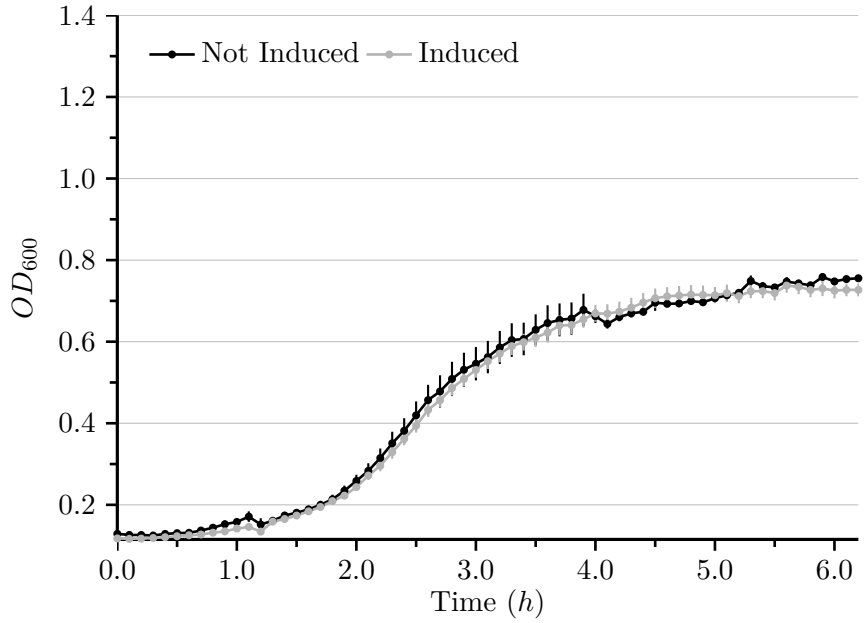


(b) Rel606

Figure 4.3: Induction of potential toxin plasmid_Toxin from *B. pseudomallei*, under the P_{BAD} promoter Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (black) or without (gray) 0.2% arabinose. Constructs were tested in (a) *E. coli* strain MG1655 and (b) *E. coli* strain Rel606. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.

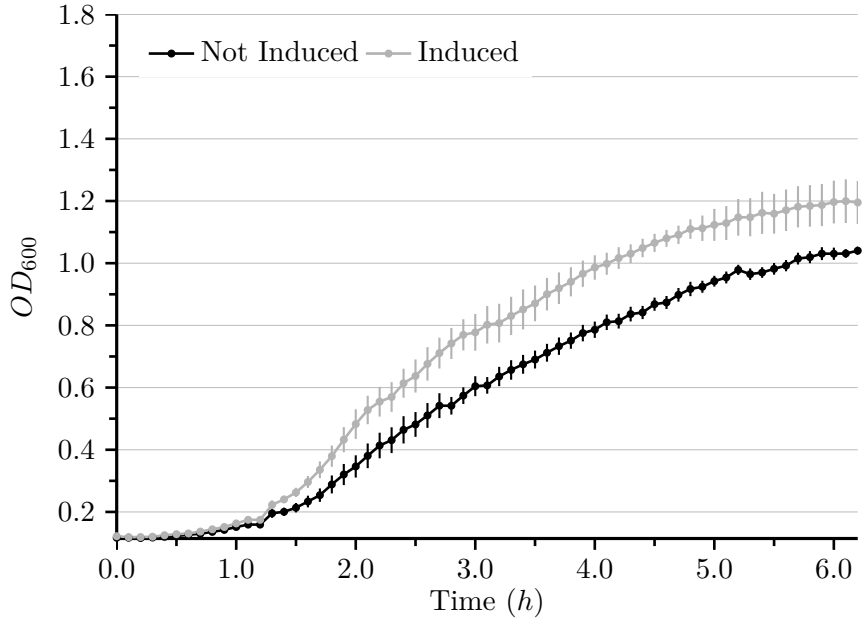


(a) MG1655

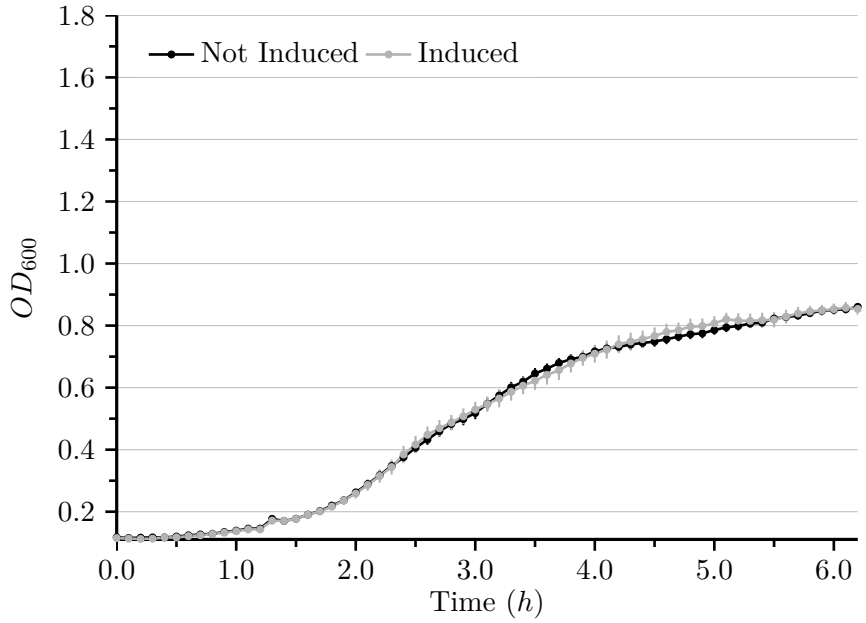


(b) Rel606

Figure 4.4: Induction of potential toxin plasmid_Toxin from *X. campestri*, under the P_{BAD} promoter Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (black) or without (gray) 0.2% arabinose. Constructs were tested in (a) *E. coli* strain MG1655 and (b) *E. coli* strain Rel606. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.

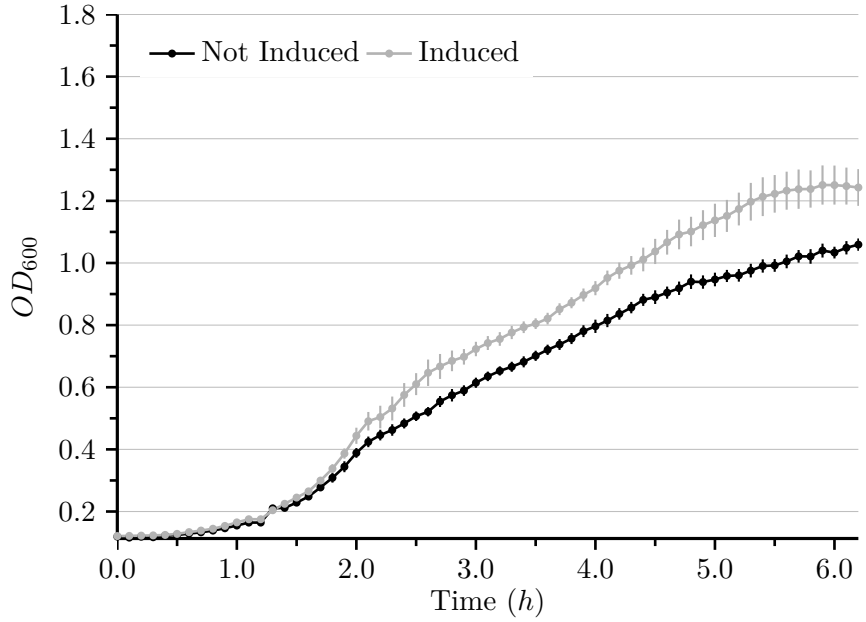


(a) MG1655

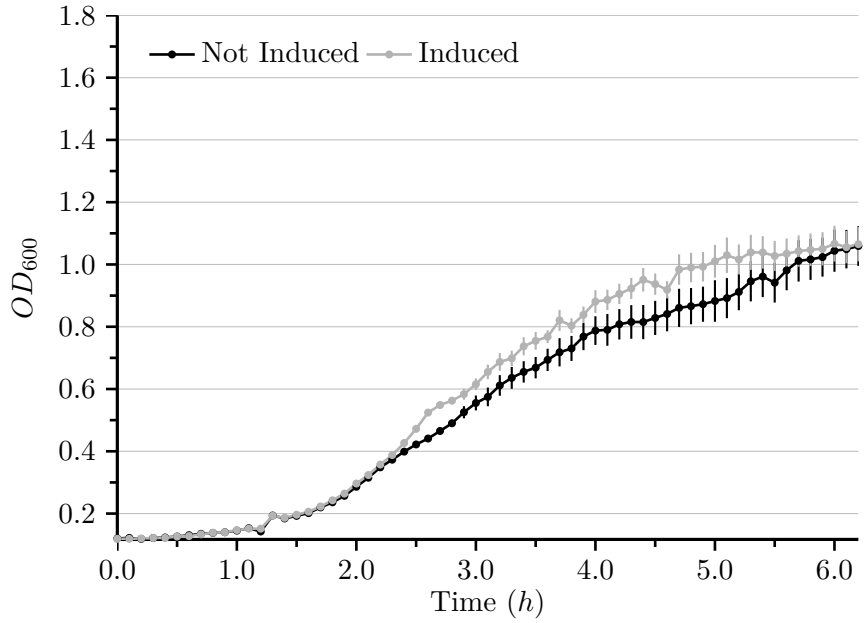


(b) Rel606

Figure 4.5: Induction of potential toxin plasmid_Toxin from *E. coli* OK137, under the P_{BAD} promoter Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (black) or without (gray) 0.2% arabinose. Constructs were tested in (a) *E. coli* strain MG1655 and (b) *E. coli* strain Rel606. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.

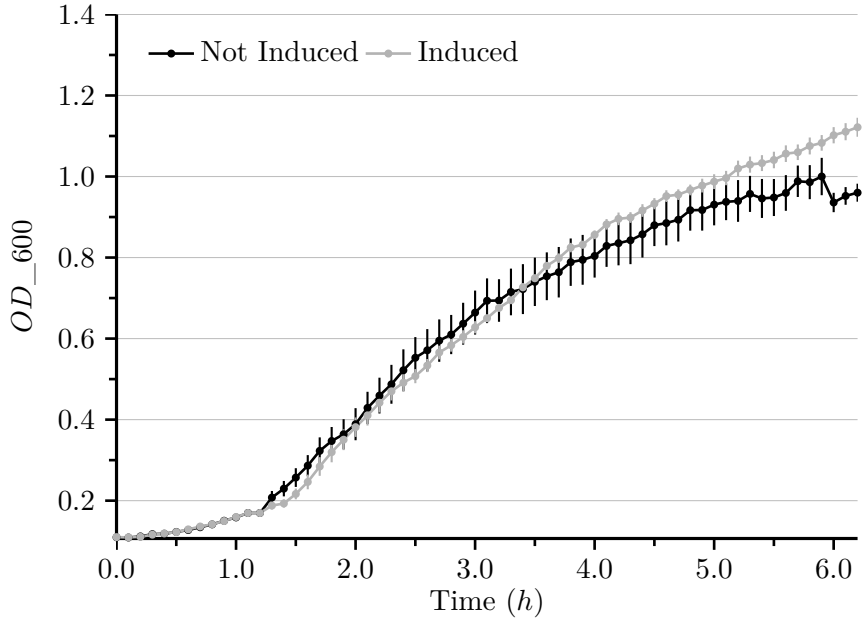


(a) MG1655

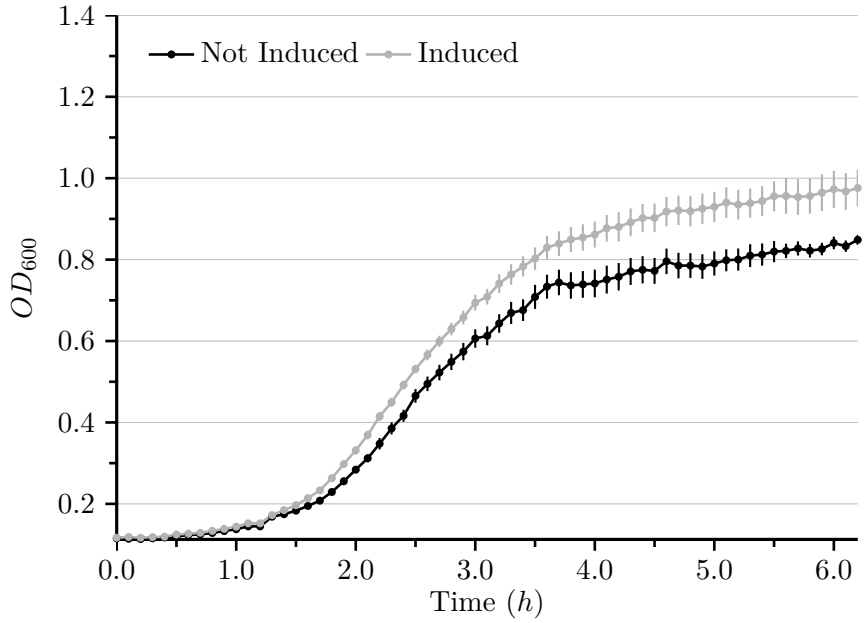


(b) Rel606

Figure 4.6: Induction of potential toxin plasmid_Toxin from *P. aeruginosa* plasmid pMATVIM-7, under the P_{BAD} promoter Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (black) or without (gray) 0.2% arabinose. Constructs were tested in (a) *E. coli* strain MG1655 and (b) *E. coli* strain Rel606. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.



(a) MG1655



(b) Rel606

Figure 4.7: Growth of *E. coli* bearing pBAD33, with and without arabinose treatment Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (black) or without (gray) 0.2% arabinose. Constructs were tested in (a) *E. coli* strain MG1655 and (b) *E. coli* strain Rel606. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.

cultures with the constructs were grown to mid log phase and induced with arabinose. Cells were harvested at late log phase, and RNA was extracted. The RNA was treated with DNase and tested with construct-specific primers to confirm there was no residual DNA (Figure 4.8(H) and Figure 4.9(H)). The primers were also tested on Rel606 DNA, to show the primers did not amplify a non-target sequence (Figure 4.8(G) and Figure 4.9(G)). The clean RNA was used to generate cDNA via reverse transcription, in turn tested for presence of the PT genes with PCR (Figure 4.8(D) and Figure 4.9(D)). Primers for the constitutively expressed glutathionine S-transferase were used as a positive control to ensure quality of the cDNA (Figure 4.8(B) and Figure 4.9(D)).

A transcript of expected size was detected amongst the RNAs of bacteria with all of the pBAD_PT constructs, the synthesized and the cloned (Figure 4.8(D) and Figure 4.9(D)). Transcription does not ensure translation into functional protein, but does suggest that the test expression system is functional.

The P_{BAD} promoter was used here because it is highly repressible, ideal for working with potentially toxic proteins. But it is about ten times less inducible than the Lac promoter (Guzman *et al.* 1995). It could be that higher levels of protein expression would show an affect on growth. Other type I toxins have been shown to have different effects at different levels of expression. The type I toxin Fst targets the cell membrane and is lethal at high levels, but also affects chromosomal segregation and cell division at low levels (Patel and Weaver 2006).

4.1.2 Heterologous expression of candidate type I toxins on high copy number vectors

Toxin genes from the two cloned systems, PT_EC and PT_PA were also cloned into the high copy number plasmid pGEM with a T7 RBS. The *E. coli*-derived PT, PT_EC, was deemed most likely to be functional in the test strains. The plasmid-borne *P. aeruginosa* PT was also a good candidate, as plasmid-borne TA toxins are more likely to be toxic than those found on chromosomes (Mine *et al.* 2009). The pGEM T-easy vector has two promoters, a T7 promoter upstream of the MCS and a P_{lac} promoter on the opposite strand downstream. The cloning strain DH5α does not contain a T7 polymerase, so genes downstream of the T7 promoter are not transcribed. The constructs were originally tested in Rel606, an *E. coli* B strain and

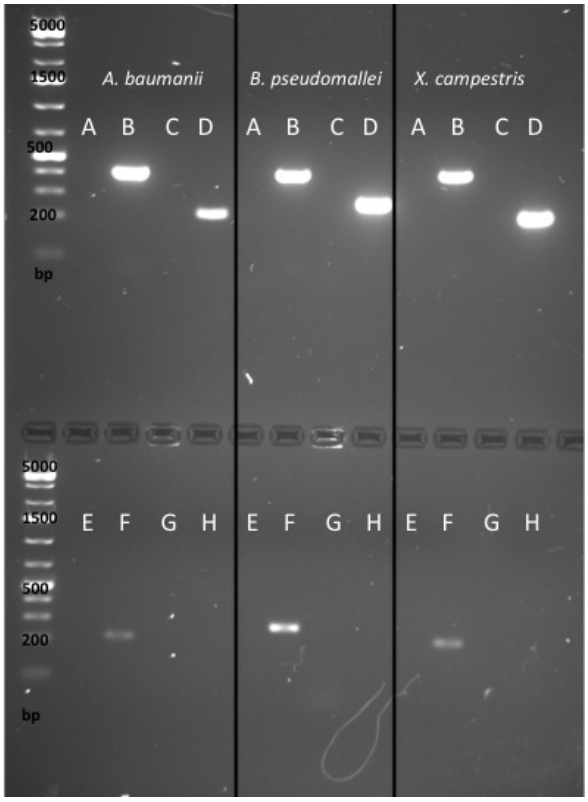


Figure 4.8: Transcription of synthesized candidate plasmid_Toxin genes in *E. coli* strain Rel606. Bacteria containing PT genes synthesized under the P_{BAD} promoter were cultured to mid-log phase and induced with 0.2% arabinose. cDNA was synthesized from total RNA preps, and tested for PT gene transcription. PCR from cDNA: A) water control B) positive control (glutathionine S-transferase, constitutively expressed in *E. coli*) C) negative control (PT primers on Rel606 RNA). D) PT primers on R6 pBD_PT cDNA. PCR from total RNA preps to detect contaminating DNA: E) water control F) positive control (pBD_PT DNA added to RNA, tested with PT primers) G) Rel606 DNA H) PT primers on R6 pBD_PT RNA.

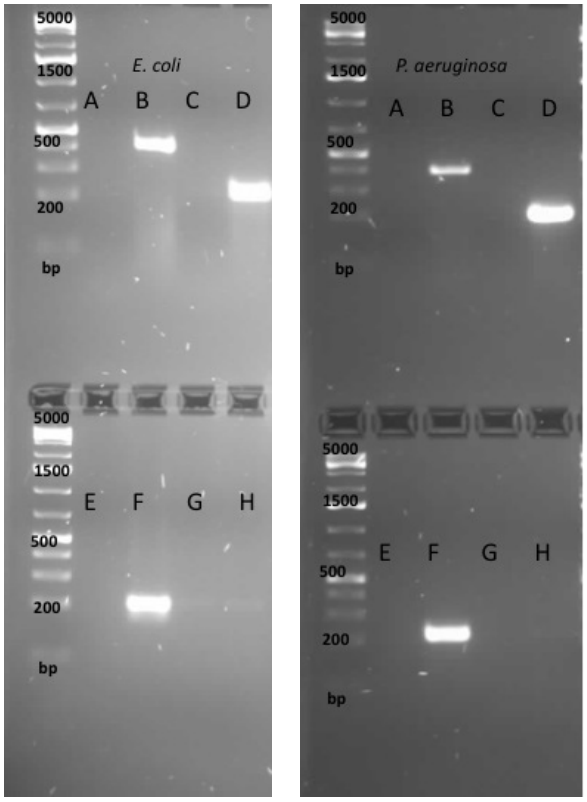


Figure 4.9: Transcription of synthesized candidate plasmid_Toxin genes in *E. coli* strain Rel606. Bacteria containing PT genes cloned under the P_{BAD} promoter were cultured to mid-log phase and induced with 0.2% arabinose. cDNA was synthesized from total RNA preps, and tested for PT gene transcription. PCR from cDNA: A) water control B) positive control (glutathionine S-transferase, constitutively expressed in *E. coli*) C) negative control (PT primers on Rel606 RNA). D) PT primers on R6 pBD_PT cDNA. PCR from total RNA preps to detect contaminating DNA: E) water control F) positive control (pBD_PT DNA added to RNA, tested with PT primers) G) Rel 606 DNA H) PT primers on R6 pBD_PT RNA.

BL21(DE3), an *E. coli* B strain with an IPTG inducible T7 polymerase to allow transcription of the constructs downstream of T7 promoter. The PT_PA construct downstream of P_{lac}, pGEM_PAL, was induced in Rel606, and the reverse orientation plasmids, pGEM_ECT and pGEM_PAT, were tested in BL21(DE3).

Of ten clones of PT_EC inserted into pGEM, all were aligned to produce the correct mRNA from the T7 promoter but not the P_{lac} promoter. As the cloning of the PCR products introduces no orientation bias, this suggested that PT_EC was toxic under the presumably leaky P_{lac} promoter. In contrast, both orientations were obtained for PT_PA.

The constructs proved to be very unstable in these strains. Colonies containing the plasmids grew slowly. Plasmid-free segregants were commonly found following liquid culture. Presumably this is a result of the higher growth rate of plasmid-free cells coupled with the depletion of ampicillin by the secreted b-lactamases of plasmid-bearing cells. This has been seen in this laboratory with other toxin-expressing, ampicillin-selected plasmids (personal observation). This was not relevant to previous experiments with PT, which selected for plasmid-bearing bacteria with other antibiotics, either chloramphenicol (pBD_PT) or chloramphenicol and kanamycin (pHS_PT).

By the time a liquid culture reached saturation, the majority of cells tended to be plasmid-free, so the affect of toxin induction was tested on solid media. Colonies were picked from plates containing LB medium supplemented with ampicillin and restreaked onto solid RM media with glucose or glycerol, with or without IPTG. For strain R6 pGEM_PAL, with PT_PA under the Lac promoter, induction with IPTG did not result in a significant decrease in growth (Figure 4.10). The strains BL pGEM_ECT and pGEM_PAT, with PT_EC and PT_PA under the T7 promoter, still did not consistently grow on solid media. Some colonies, upon restreaking onto new plates, grew well and some did not grow at all, even between plates of the same medium. This was possibly due to the differences in the number of plasmid-containing cells within each colony, resulting in difficult to interpret results.

Interestingly, plasmids were uniformly maintained in the DH5α strain in liquid and solid media. Cultures of the strain containing the pGEM_PT constructs were grown to early log phase and transcription was induced with IPTG (Figure 4.12). Again, pGEM_PAL did not decrease growth in these cultures (Figure 4.11b), either before or after induction of PT_PA when compared to the untransformed control

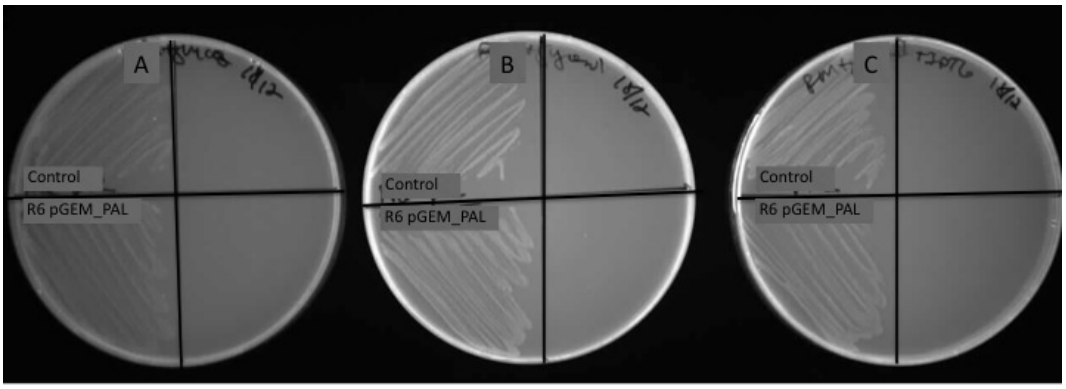
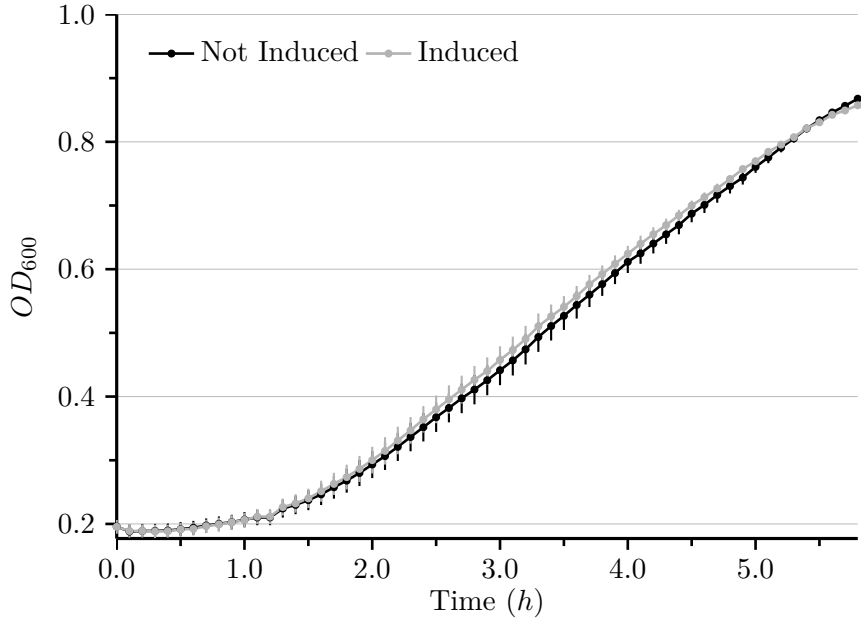
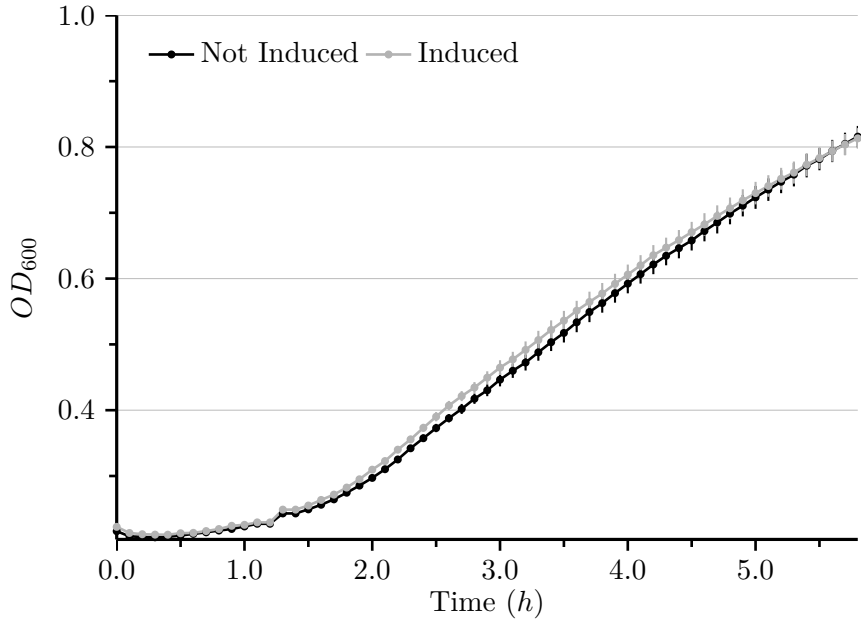


Figure 4.10: Induction of candidate toxin plasmid_Toxin from *P. aeruginosa* plasmid pMATVIM-7 under the Lac promoter in *E. coli* strain Rel606 on solid media. Colonies of Rel606 transformed with pGEM (control) and pGEM with the *P. aeruginosa* plasmid pMATVIM-7 PT ORF (R6 pGEM_PAL) were streaked onto RM media with A) glucose (repressing medium) B) glycerol (permissive medium) and C) glycerol and IPTG (inducing medium). The remaining two quadrants were unused. Representative of three replicas of the experiment.

(Figure 4.11a). Growth of DH5α with PT_PA in the reverse orientation was originally tested as a control. Transcription of the region was considered unlikely due to the lack of a T7 polymerase for transcription from the T7 promoter and RBS for transcription from the Lac promoter. Unexpectedly, bacteria containing PT_PA in the reverse orientation increased in density at a slower rate than bacteria with PT_PA downstream of P_{lac} (Figure 4.12a). A similar effect was seen with bacteria containing PT_EC in the reverse orientation (Figure 4.12b). This effect was not increased by induction of the Lac promoter with IPTG ($p>0.05$).

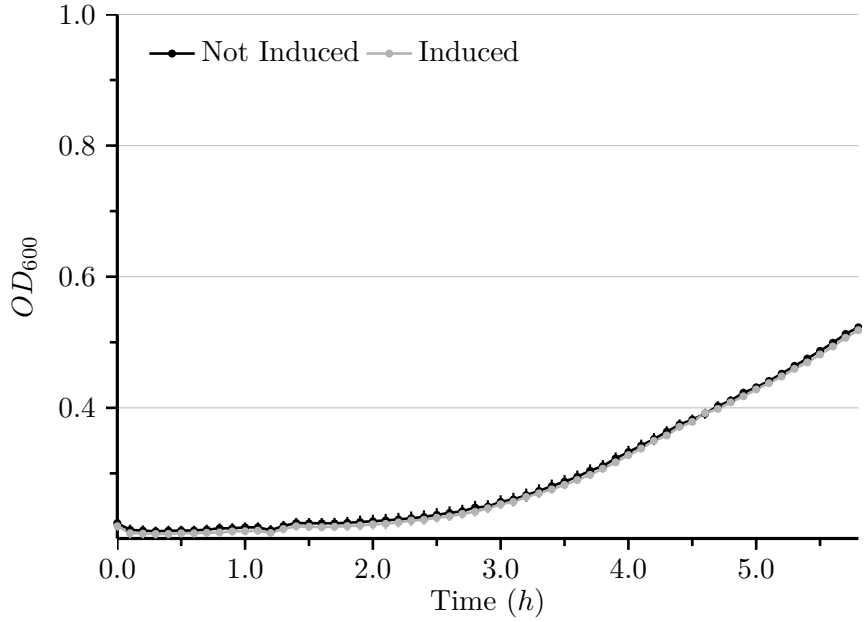


(a) untransformed cells

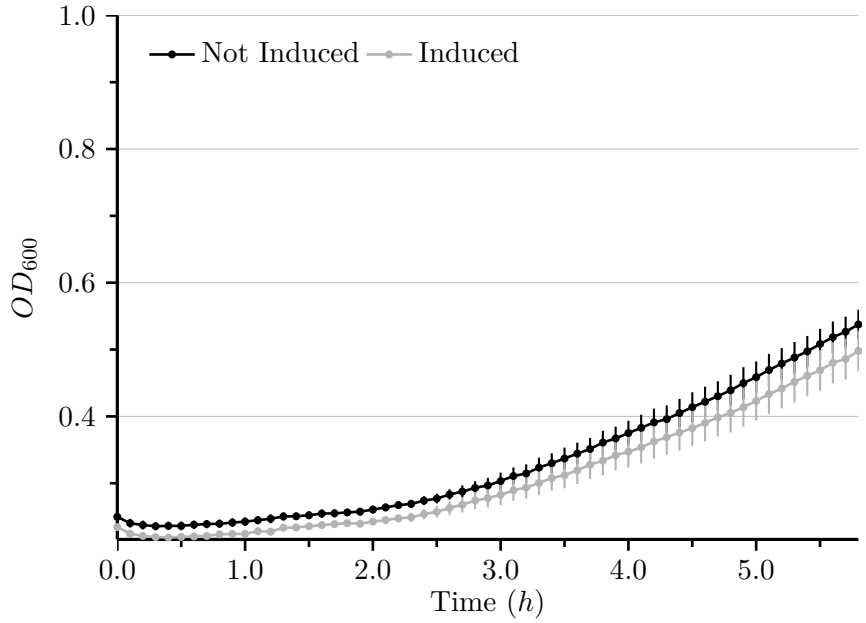


(b) PT PA

Figure 4.11: Induction of potential toxin plasmid_Toxin under the Lac promoter in *E. coli* strain DH5α, control and *P. aeruginosa* plasmid pMATVIM-7. Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (—●—) or without (—⊕—) 1 mM IPTG. PTs were cloned into pGEM vector behind the Lac promoter. (a) Control, untransformed DH5α (b) *P. aeruginosa* plasmid pMATVIM-7. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.



(a) PT PA, Reverse orientation



(b) PT EC, Reverse orientation

Figure 4.12: Induction of potential toxin plasmid_Toxin from *P. aeruginosa* plasmid pMATVIM-7 and *E. coli* OK137 in the reverse orientation under the Lac promoter in *E. coli* strain DH5α. Bacteria were cultured to 0.1-0.2 OD_{600} before induction with (—●—) or without (—⊕—) 1 mM IPTG. PTs were cloned into pGEM vector behind the Lac promoter. (a) reverse orientation (b) *E. coli* OK137, PT_EC, reverse orientation. Experiments were repeated in three independent trials, with four within trial replicas, ± SE.

4.1.3 Testing PT-ptaRNA1 operons for PSK

The entire predicted operons for PT-RNA1 loci in *E. coli* (PR1_EC) and *P. aeruginosa* plasmid pMATVIM-7 (PR1_PA), were synthesized and cloned using the vector pHSG415. Again, the *E. coli*-derived operon, PR1_EC, was assumed to be the most likely to be functional in the test strains. The PR1_PA was chosen because plasmid-borne TA systems are more likely to confer the PSK phenotype than TA systems found on chromosomes (Mine *et al.* 2009). Neither PT showed toxicity in the test strains. Toxicity is required for PSK. The operons were tested for PSK regardless, as the full sequence of the PT ORFs as expressed in their native context is uncertain and it was possible that the larger region would be functional.

Along with the predicted PT_ORF and ptaRNA1 regions, an additional 60-70 bp downstream and 222 bp upstream of the toxin were synthesized in hopes of including regulatory regions. Regulatory regions of type I TA systems differ significantly. Some type I TA toxins have long 5' UTRs where the RNA binds. This places the promoter far upstream of the ORF. Though ptaRNA1 is antisense to and believed to bind the RBS of PT (Findeiß *et al.* 2010), which is not usually associated with long 5' UTRs, the region of overlap is not always the region of regulation. In the SprA1 type I system, regulatory RNA interacts with regions of toxin mRNA outside of their region of overlap (Sayed *et al.* 2011; Sayed *et al.* 2012).

The two operons were tested for their ability to induce PSK in *E. coli* strains Rel606 and MG1655. Both strains have been fully sequenced, and did not appear to contain any ptaRNA1 loci in our screen, reducing the likelihood of anti-addiction from chromosomal antitoxins (Mine *et al.* 2009; Saavedra De Bast *et al.* 2008; Cooper and Heinemann 2005). The two operons, PR1_EC and PR1_PA, were cloned using the plasmid pHSG415. This vector is a low copy, partially deleted for partitioning genes (Caulcott *et al.* 1987) and temperature-sensitive for replication. Bacteria containing the plasmid are usually cultured at 30℃. When shifted to 42℃ the plasmid can no longer replicate along with the cells, due to mutations in the origin of replication (Hashimoto-Gotoh *et al.* 1981).

If pHSG415 contains a TA system that can confer a PSK phenotype, cells will die as at least one daughter fails to inherit a plasmid. If the plasmid does not contain such a system, plasmid-free daughters survive and accumulate. Growth is reported as the log ratio of total cells to plasmid-containing cells over time (Figures 4.13

and 4.14). In cultures containing a '+PSK' plasmid, the ratio of cells would remain 1:1 (100), as plasmid-free cells would die upon segregation. The ratio would increase in cultures with a '–PSK' plasmid, as cells continue to replicate while the plasmid does not.

The PR1 operons did not exhibit PSK in either cellular background (Figures 4.13 and 4.14). They had a pattern identical to the no-PSK control, where the number of total CFUs in the culture (as measured on LB plates) grows exponentially while the number of CFUs retaining the plasmid (as measured on a selective medium) stays the same. The ratio of bacteria without pHS_EC and pHS_PA (a measure of PSK) were not significantly different than the control pHSG415 ($p>0.05$) at four hours except for pHS_EC in Rel606, where the ratio was actually significantly ($p=0.012$) *higher* than the control.

As controls, two known PSKs were tested. One was the RM system PaeR7 and the other a type II TA system called ParDE. In MG1655, the plasmid containing PaeR7 (pTN9) demonstrated one phenotype consistent with that predicted of a PSK. That is, the culture ceased to grow at 42℃ and the ratio of plasmid-containing cells to total cells remained about 1:1. This ratio was significantly different than the control pHSG415 ($p=0.001$) at four hours. *E. coli* cultures with a plasmid carrying the ParDE system (pHS_par), however, demonstrated an intermediate phenotype, though it was still significantly different than the control ($p=0.007$). This suggests that not all plasmid less daughters died. Interestingly, PaeR7 did not cause PSK in a Rel606 background (Figure 4.14). This illustrates the difficulty in testing PSK. There may be several reasons for this. There could, for example, be a copy of the antitoxin on the chromosome. PSK may not occur were the operons incomplete as synthesized, or if regulation within the test strain was not optimal. Rel606 does not produce a Lon protease, which would affect the PSK of type II TA systems as it degrades many protein antitoxins, but would not be expected to affect the PSK of PaeR7 (a RM system) or type I TA systems. Finally, appropriate expression of the toxin would be necessary for a PSK effect. Both operons were tested for toxin transcription.

Transcription of candidate toxins from PT-ptaRNA1 operons

Transcription of the PT ORF was tested for both constructs, using RT-PCR. As with the pBAD_PT constructs, bacteria were cultured to late log phase before collection

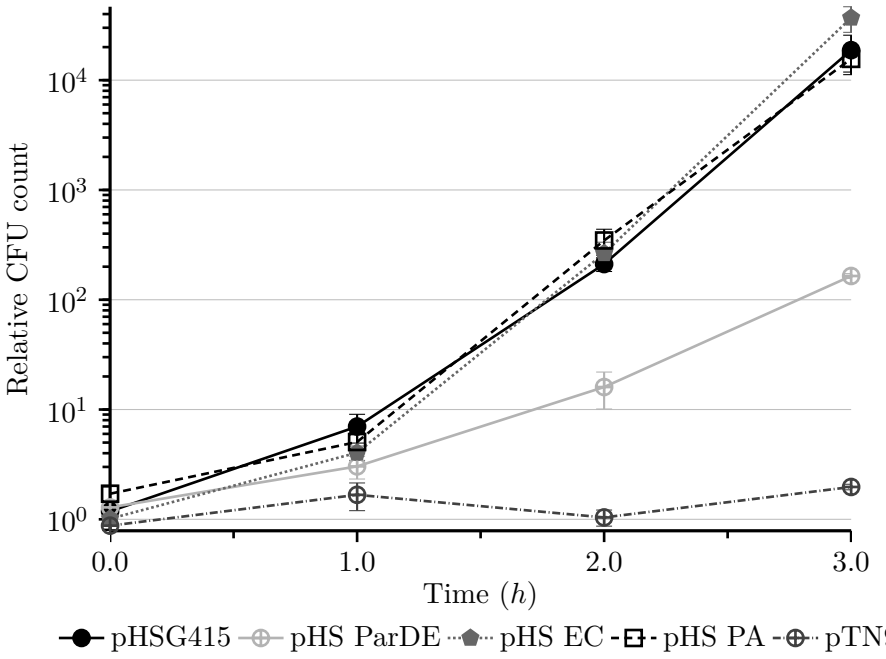


Figure 4.13: Effect of potential PSK-inducing TA systems on bacterial growth under conditions of forced plasmid loss in *E. coli* strain MG1655. Bacteria containing potential PSK-inducing TA systems on temperature sensitive plasmids were cultured at 42℃to force plasmid loss. To obtain relative CFU count, number of total live cells (plated on LB) was divided by cells containing the plasmid (plated on antibiotics). Operons inducing PSK would be expected to kill cells losing the plasmid, giving a relative CFU count of ~1. (—●—) temperature sensitive (TS) plasmid pHSG415, no PSK. (—⊕—) TS plasmid pHS_ParDE with known PSK ParDE. (···⬟···) TS plasmid pHS_EC with pTARNAi operon from *E. coli*, PR1_EC (-⊡-) TS plasmid pHS_PA with pTARNAi operon from *P. aeruginosa* plasmid pMATVIM-7, PR1_PA. (-⊕-) TS plasmid with known PSK, PaeR7. Experiments were repeated in three independent trials, with two within trial replicas, ± SE.

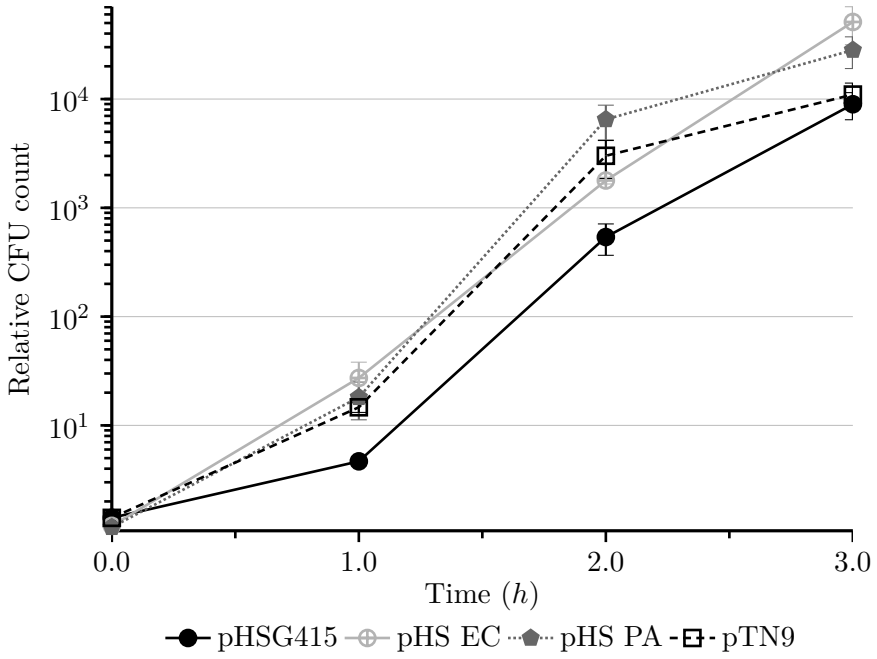


Figure 4.14: Effect of potential PSK-inducing operons on bacterial growth under conditions of forced plasmid loss in *E. coli* strain Rel606. Bacteria containing potential PSK-inducing TA systems on temperature sensitive plasmids were cultured at 42℃to force plasmid loss. To obtain relative CFU count, number of total live cells (plated on LB) was divided by cells containing the plasmid (plated on antibiotics). Operons inducing PSK would be expected to kill cells losing the plasmid, giving a relative CFU count of ~1. (—●—) temperature sensitive (TS) plasmid pHSG415, no PSK. (—⊕—) TS plasmid pHS_ParDE with known PSK ParDE. (···⬟···) TS plasmid pHS_EC with pTARNAi operon from *E. coli*, PR1_EC (-⊡-) TS plasmid pHS_PA with pTARNAi operon from *P. aeruginosa* plasmid pMATVIM-7, PR1_PA. (-⊕-) TS plasmid with known PSK, PaeR7. Experiments were repeated in three independent trials, with two within trial replicas, ± SE..

and extraction of RNA. Reverse transcription was used to generate cDNA from the RNA. Controls included testing RNA for residual DNA (Figure 4.15(H)), testing for non-target amplification using DNA from the parental strain (Figure 4.15(G)), and amplifying cDNA from constitutively transcribed glutathionine S-transferase as a positive control to show integrity of the cDNA (Figure 4.15(B)). The cDNA was tested for presence of PT genes from the operon with PCR (Figure 4.15(D)).

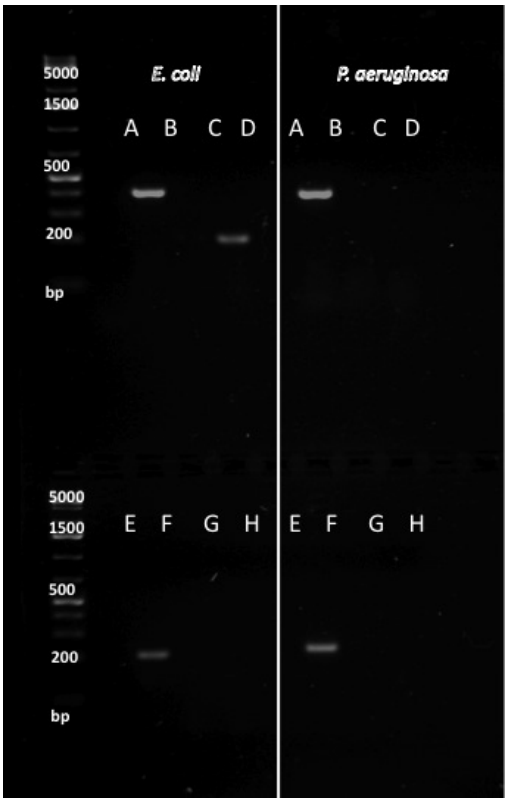


Figure 4.15: Transcription of candidate toxins from PT-RNA1 operons in different bacterial species, expressed in *E. coli* strain Rel606. *E. coli* containing synthesized ptaRNA1 operons on plasmid pHSG415 were harvested at mid log phase. cDNA was synthesized from total RNA preps, and tested for PT gene transcription. PCR from cDNA- A) water control B) positive control (glutathionine S-transferase, constitutively expressed in *E. coli*) C) negative control (PT primers on Rel606 RNA). D) PT primers on R6 pHS_PT cDNA. PCR from total RNA preps to detect contaminating DNA- E) water control F) positive control (pHS_PT DNA added to RNA, tested with pHS_PT primers) G) Rel606 DNA H) PT primers on R6 pBD_PT RNA.

The RT-PCR results showed transcription of the *E. coli* but not the *P. aeruginosa* PT (Figure 4.15(D)). This would suggest that, at least for PR1_EC, sufficient regions of the operon had been synthesized for transcription of the protein from its native regulatory system. The meaning behind the negative result for PR1_PA is more difficult to interpret. The full sequence length of the PR1_PA operon, necessary for function, may differ from the candidate *E. coli* TA system, the PT-PA may be more tightly repressed by the antitoxin RNA making detection difficult, or there are too many cis-acting differences to be expressed heterologously in *E. coli*.

4.2 Discussion

Like many newly identified type I TA systems, PT-RNA1 was discovered computationally. Interestingly, chromosomal homologues of PT-RNA1 were found in a wide range of bacterial species, including beta and gamma proteobacteria, distinguishing it from other type I TA systems. Furthermore, additional examples of it could be found on mobile elements or on the chromosome of one strain but not close relatives (Findeiß *et al.* 2010). This type of distribution is indicative of dissemination by HGT which is more common in type II TA systems. This made ptaRNA1 an interesting candidate to test experimentally, giving us a chance to examine whether the operon is indeed a type I TA system, and, if so, if the distribution of this novel system is reflected in an ability to exhibit PSK. If it did, then it would likely also be successful reproducing by horizontal gene transfer. Other than a Northern blot showing expression of the RNA in exponential and stationary phase (Findeiß *et al.* 2010), no laboratory confirmation of the system has been reported.

4.2.1 Plasmid_Toxin expression did not decrease *E. coli* growth

The toxicity of five proposed PT ORFs from various species was tested under two expression systems. Under P_{BAD}, on a mid-copy number vector, induction of the ORFs did not reduce growth of *E. coli*, despite the region being transcribed. This vector and expression system has been used previously to test toxins from potential TA systems identified computationally (Leplae *et al.* 2011; Fozo *et al.* 2010). Though *E. coli* is not the original genetic background for four of the five loci, in most cases other type I toxins, even those only found in Gram-positive bacteria, have been shown to function in *E. coli* (Fozo *et al.* 2010; Weaver *et al.* 2009). There is still

a chance, though, that the effect of the protein is species specific. In this case, we might have expected that PT_EC, at least, would have proven toxic.

Toxicity was also tested at higher levels of expression from the pGEM T-easy vector system. The high copy, high expression vectors proved to be unstable in cultures of *E. coli* B strains Rel606 and BL21, though tests on solid media showed that the transcription of PT_PA under the Lac promoter did not further decrease growth in Rel606. Tests in the cloning strain DH5α showed that growth of cultures containing constructs with PT downstream of the T7 promoter as greatly reduced. As DH5α does not produce the T7 polymerase, it is unlikely to be due to expression from this promoter. It is possible that something toxic is being produced from the Lac promoter on the opposite strand despite the lack of RBS. Though, as before, there is no evidence that induction of this promoter further reduces growth. So while the pGEM constructs appear to have an affect on bacterial growth, it is difficult to say from the current evidence what the mechanism is.

There are some things that may be preventing a toxic effect from the PT constructs. The antitoxin might still be expressed from these constructs. The PtaRNA1 promoter region appears to be embedded within the PT gene. All constructs synthesized showed similar predicted promoters as seen in (Findeiß *et al.* 2010). In the native operon, the antisense RNA is believed to extend across the RBS (Figure 4.1). The embedded promoter could still be determining transcription of an antisense RNA, which would include the first six base pairs of the ptaRNA1 RNA then extend across the synthesized T7 RBS. It is possible that this construct could bind and interfere with PT translation. While transcription levels from this promoter in *E. coli* are unknown, it would have to be high enough to counteract induction of the gene from the P_{BAD} promoter.

It may be that the regions synthesized here did not incorporate all of the functional protein. ORF boundaries were determined computationally, using ORF predicting software. These algorithms may fail to always accurately predict ORFs. The PT ORFs therefore may need to be experimentally defined. In the future, instead of placing a predicted ORF behind a T7 RBS, a larger region could be cloned and placed downstream of a promoter and translated from its native RBS. If the larger region produced a toxin, mutational analysis could then be done to determine the size of the ORF. Transcriptomic techniques could also be used to identify the precise sequences of the toxin and antitoxin RNAs.

It could also be that PT can be toxic but that these particular ORFs do not encode a toxin. Even within well-studied TA system families, 'toxins' from a given loci are not always toxic. Chromosomal homologues are often no longer toxic (Pedersen and Gerdes 1999; Mine *et al.* 2009). This is true of those that appear to be in a process of inactivation through drift, as well as those known to still function in the cell, such as with HipA7 (Korch *et al.* 2003). All but one of the loci tested here were located on chromosomes. Toxin genes on mobile elements are more likely to produce proteins that are toxic (Mine *et al.* 2009).

These issues aside, the lack of toxicity from induce PT constructs could simply mean that they are not toxins. Indeed, though the effect varied depending on strain and construct, the proteins seemed to increase bacterial growth when induced. This was not entirely attributable to an increase in sugar (arabinose) in the media, as the strain unaffected by sugar in the control was the same strain that had the strongest increase in growth in response to PT induction. Growth in batch culture was the only phenotype of gene induction analyzed here. It may proove that PT-ptaRNA1 is not a type I TA system at all, but encodes a different RNA-antitoxin regulated function.

4.2.2 PT-ptaRNA1 operons did not exhibit PSK

Two PT-ptaRNA1 loci were tested in their native operon for the ability to create a PSK phenotype when expressed in *E. coli*. The two candidate TAs were from *E. coli* and the *P. aeruginosa* plasmid pMATVIM-7. Though neither PT showed toxicity, which is required for PSK, the operons were tested in the chance that expression in their native genetic context would make functional proteins. The constructs were cloned using a plasmid with temperature-sensitive replication, where plasmid loss can be forced.

Neither of the two constructs tested positive for PSK. As with the toxicity tests, the region tested was previously defined computationally, meaning that there may be regions crucial for PSK but were inadvertently not included. The full region could be determined using techniques such as RACE, which involves sequencing entire transcripts from smaller known regions. Identifying the full sequence of the native PT mRNA would allow us to narrow down on the regions necessary for PSK. But, while experimental determination of operon region may be helpful, RT-PCR data

showed transcription of PT-EC, indicating that at least that region is sufficient for PT ORF transcription. The toxin gene in *P. aeruginosa* operon was not transcribed, possibly explaining the lack of PSK.

PSK is a more complex phenotype than toxicity alone, with exhibition requiring toxicity within the host system as well as appropriate regulation of the toxin and antitoxin. If the PT ORFs do indeed prove to not encode a toxin for *E. coli*, one would not expect the PT-ptaRNA1 operons to exhibit PSK. If toxicity is only something that occurs at high levels, not low levels, as with Fst, PSK would be contingent on toxin expression level. To assure proper regulation of the genes within the operon, PSK is best tested within the native host. Interference from the resident locus would need to be negated by knocking out the locus or curing the host of the resident mobile element containing it. The *P. aeruginosa* locus tested here would be an ideal candidate for testing within its host. The locus is plasmid borne, so it is the most likely to induce PSK. Plasmid-free cells could be infected with the locus on a temperature sensitive or vertically unstable plasmid, and monitored accordingly. The lack of PSK seen in PR1_EC, even when the proposed toxin is being transcribed under its native promoter, suggests that the operons are unlikely to exhibit PSK in their native hosts.

While further experiments can be conducted to confirm the results presented here, some conclusions can be drawn by taking these results at face value. The PT-PtaRNA1 was first presented as a potential type I TA system. PT-PtaRNA1 has a distribution indicative of that caused by HGT and retains the operon structure typical of a type I TA system, but does not appear to have the associated biochemical features exhibited by PSK systems. According to the structural definition provided in Chapter 1, a TA system should be composed of a toxin and an antitoxin, be organized into an operon, and have antitoxin-mediated regulation of the toxin. The lack of toxicity from PT ORFs would suggest that PT-PtaRNA1 is not a TA system and also not a PSK system. Our original interest in the PT-ptaRNA1 system was due to its unique distribution when compared to other type I TA systems. I hypothesized that this system was distributed more like type II TA systems, and had a high chance to exhibit PSK. As PT-ptaRNA1 did not appear to be a type I system from these tests, it is difficult to derive any conclusions about it in relation to other TA systems.

PT-ptaRNA1 may have a non-PSK function that selects for its presence on mobile elements. The Kis-Kid type II TA system has been shown to act primarily

pre-segregationally rather than post-segregationally, though it can alleviate bacterial growth under conditions of forced loss (Pimentel *et al.* 2005). The bacteriostatic toxin Kid becomes activated before plasmid-loss, when the plasmid is at low copy numbers. Kid degrades selected mRNAs (Pimentel *et al.* 2005; Pimentel *et al.* 2014), resulting in reduction of cell division and depression of plasmid replication, allowing the plasmid to rapidly increase in copy number (Pimentel *et al.* 2014).

PT-ptRNA1 retains the structural similarity, as identified using bioinformatics, to a type I TA system but has an altered biochemical profile. Bioinformatics is an important tool in genetics, allowing us to sift through increasingly large databases of sequencing information. There are limitations, however, in what can be inferred without experimental work to verify computationally derived hypotheses.

Chapter 5

Engineering expression systems for exhibition of PSK by a given toxin and antitoxin, using barstar and barnase as model proteins

In the second part of this thesis, I use toxin and antitoxin gene pair barnase and barstar from the chromosome of *B. amyloliquefaciens* as a model system to explore the conditions necessary for PSK. Barnase is a small (12.4 kDA) extracellular RNase bound in a one to one ratio by its inhibitor, barstar (10.2 kDA) (Hartley and Smeaton 1973; Hartley 1988). Condon and Putzer (2002) found some evidence the genes were distributed across chromosomes in a manner inconsistent with purely vertical inheritance. This included only the barstar gene in *B. subtilis*, both genes in *Clostridium acetobutylicum* and *Yersinia pestis*, and an N-terminal truncated barnase followed by barstar in two independent sequences of *Neisseria meningitides* (Condon and Putzer 2002). One research group also suggested that plasmids with barnase and barstar were easy to work with due to their stability in the absence of an external selection, such as an antibiotic, for at least 36 generations (Ul'yanova *et al.* 2007). Nonetheless, they are not known to exhibit PSK.

In this chapter I ask the question: given that genes can move into new cellular environments, potentially free of their native regulatory networks, what characteristics are necessary for toxin and antitoxin pairs to confer a PSK-like phenotype? Barstar and barnase were placed in a new cellular (*E. coli*) and genetic (plasmid) context. A range of expression systems with which to test barstar and barnase were developed in *E. coli* using inducible and constitutive promoters. The aim was to develop a

suite of suitable vectors that enabled positive and negative regulation of barstar and barnase expression, at varying proportions to one another, for the eventual purpose of testing the genes for PSK activity. The levels of gene expression in relation to protein stability necessary for barstar and barnase to act as a PSK in a manner similar to type II TA systems were then explored. I draw some preliminary generalizations for all TA systems based on the results.

5.1 Results

5.1.1 Expression systems with barnase under the P_{tac} promoter

A clone of the barnase gene was gifted to us on the pUC-derived plasmid pMT416 (Hartley 1988). It was renamed pBB01 for the purpose of this work. On this plasmid barnase was fused to the *phoA* signal sequence for secretion to the periplasmic space via the *E. coli* secB-secA translocation pathway (Michaelis *et al.* 1983; Li *et al.* 1988). Expression was under control of P_{tac}(Maloney and Rotman 1973) (Figure 5.1), allowing it to be induced with IPTG (Hartley 1988). Barstar is downstream of its native promoter from *B. amyloliquefaciens*, which is expressed constitutively in *E. coli* (Hartley 1988).

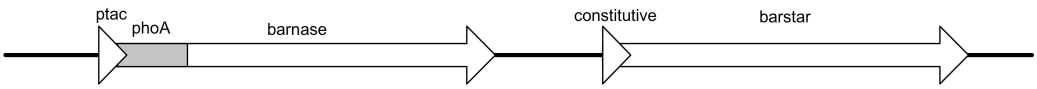


Figure 5.1: Organization of barstar and barnase genes on plasmid pBB01. The RNase barnase is under control of P_{tac}, which can be induced with IPTG at the transcriptional level. The signal sequence *phoA* targets the protein for secretion into the periplasmic space of *E. coli*. Downstream, the intracellular inhibitor barstar is under a constitutive promoter. Genes are on a high-copy number pUC-derived backbone.

Toxin and antitoxin activity of this system was tested in CSH100, a well-characterized *E. coli* strain with the $lacI^Q$ gene on an F plasmid for tighter inhibition of P_{tac} when not induced. Attempts at transforming non- $lacI^Q$ *E. coli* strains with this plasmid were unsuccessful, indicating that leaky expression from the P_{tac} promoter is lethal. CSH100 transformed with pBB01 (CSBB01) were treated with IPTG to induce bar-

nase expression and monitored for growth, measured as optical density (OD_{600}) of the culture (Figure 5.2) as well as ability to form colonies (CFUs) on petri dishes with ampicillin (Figure 5.3). Controls involved bacteria transformed with pP01, a pUC-derived plasmid with a P_{tac} - *lacZ* fragment (Figure 5.2).

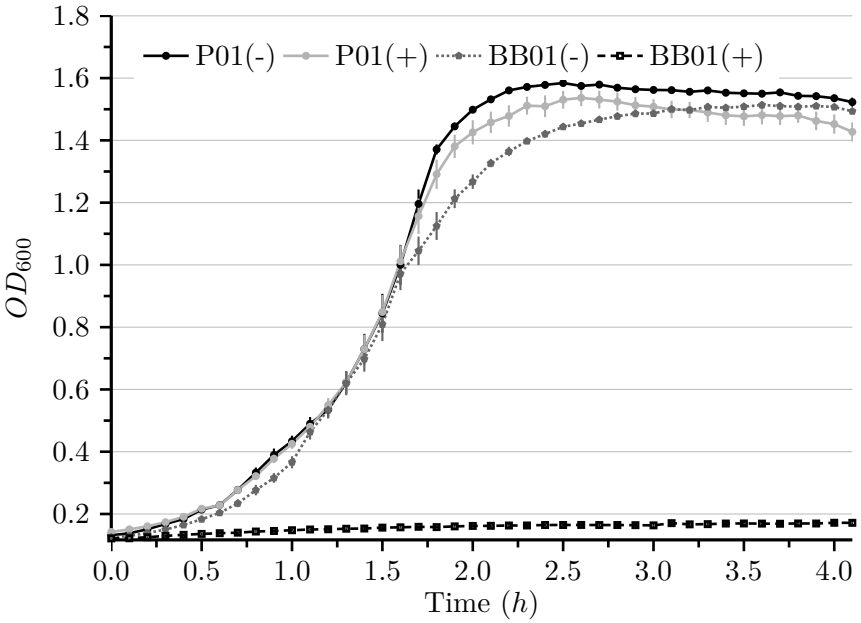


Figure 5.2: Optical density of *E. coli* culture expressing barnase and barstar under the P_{tac} promoter. Cultures were grown to density of ~0.100 OD_{600}, then split before treatment (time zero) with or without 1.0 mM IPTG. (–●–) CSP01(–), uninduced. (–⊕–) CSP01(+), *lacZ* induced. (···⬟···) CSBB01(–), uninduced. (-⊟-) CSBB01(+), barnase induced. Time is given as hours after induction. Measurements were the result of three independent trials with four replicas per trial, ± SE.

The addition of IPTG and assumed subsequent expression from the P_{tac} promoter resulted in only a small decrease in the viability of CSP01 cultures. Cultures of CSBB01, on the other hand, did not increase in optical density once IPTG was added (Figure 5.2), significantly different than the uninduced control at four hours ($p \leqslant 0.001$). This is similar to what has been seen elsewhere (Hartley 1988; Ramos *et al.* 2005). However, optical density does not measure the number of viable cells in the culture. To distinguish between total and viable bacteria, CFUs were enumerated from samples taken at one-hour time points after induction. CFU count was also

significantly reduced over this same period (p=0.001 at four hours, (Figure 5.3)), suggesting that the barnase gene was being expressed to a toxic protein in CSBB01.

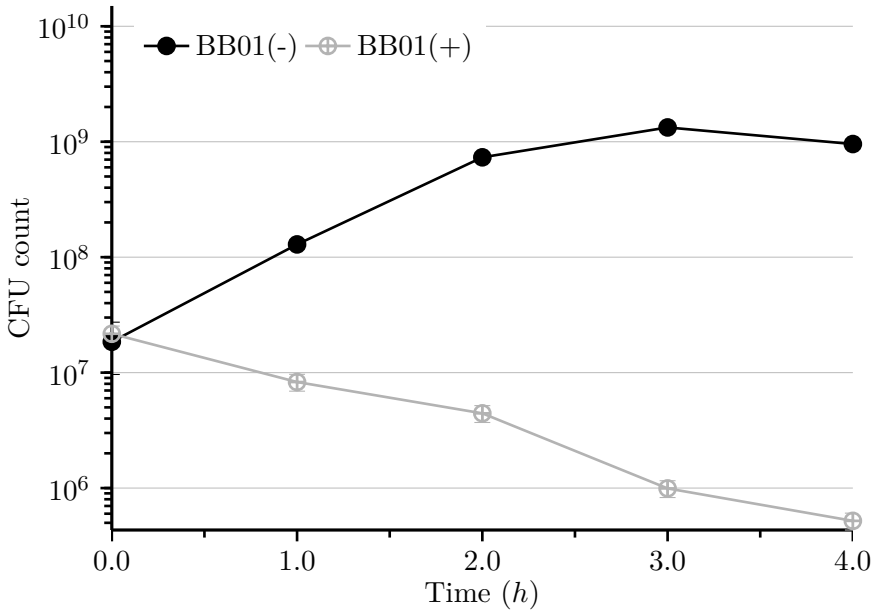


Figure 5.3: CFUs of *E. coli* culture expressing barnase and barstar under the $\mathbf{P_{tac}}$ promoter. Cultures were grown to a density of ~0.100 OD_{600}, then split before treatment (time zero) with or without 1.0 mM IPTG. Samples were diluted and plated on selectable media to enumerate CFUs. (–●–) CSBB01(–), uninduced. (–⊕–) CSBB01(+), barnase induced. Measurements were the result of three independent trials with two replicas per trial, ± SE.

Barnase reinduction

The growth of bacterial cultures was measured in response to adding and removing inducers of barnase expression (Figure 5.4). Cultures of CSBB01 were alternately induced, inducer removed by dilution through washing, and inducer re-introduced over 24 hours. Growth of the culture was measured by CFU count.

As before, cultures responded to barnase induction by a reduction in CFU count (Figure 5.4). Cultures that were subsequently washed of IPTG saw an increase in CFU count over three hours, which continued up to 24 hours to levels similar to the non-induced culture. To show that the functional toxic protein could still be expressed in these cells, some cultures were re-exposed to inducer three hours after

washing. These cultures exhibited a decrease in CFU count again, down to levels similar to that seen in cultures grown continually in inducer.

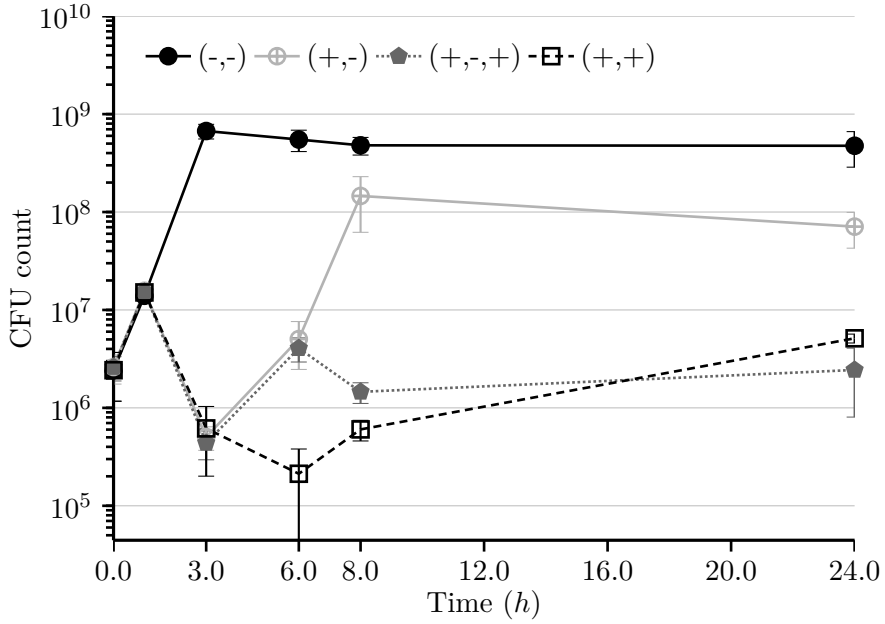


Figure 5.4: Resuscitation of *E. coli* growth after loss of barnase inducer, with barnase on a high copy number plasmid. Cultures of CSBB05 were sampled and plated on antibiotics to determine the number of CFUs at time intervals indicated. At one hour, all but (–●–) (–,–) were treated with 1.0mM IPTG to induce barnase expression. All samples were washed at T3 and resuspended in appropriate medium. (–⊕–) (+,–) Cells resuspended in –IPTG medium. (-⊟-) Cells resuspended in –IPTG medium, then re-exposed to IPTG at T6. (···⬟···) (+,+) Cells resuspended in +IPTG medium. Measurements are the result of three independent trials, with two replicas per trial, ± SE.

It was common during the re-induction assay for bacteria free of the pBB01 plasmid to arise and multiply in the cultures when subjected to the inducer for long periods of time, as in this experiment. The pBB01 plasmid is selectively maintained in the culture with ampicillin. Once ampicillin is broken down in the medium by β-lactamases secreted by resistant bacteria, plasmid-free cells can quickly take over the culture. CFU counts were performed on ampicillin plates, so any plasmid-free cells in culture were not enumerated but could have confounding effects during culture by outcompeting slower growing bacteria expressing barnase.

To alleviate this, the reinduction assay was repeated with strain CSH100 transformed with a new construct, pBB03 (Figure 5.5). The pBB01-sourced fragment of barnase under P_{tac} and native barstar gene, was cloned into a medium copy vector (~20 per cell) with a pBR322 backbone. Along with being lower copy number, the vector confers kanamycin resistance. The pBB03 vector also allowed us to test expression of barnase on a low-copy number vector compared to a high copy number vector.

CSBB03 behaved similarly to CSBB01 through an induction and reinduction of barnase cycle. The effect of the inducer was reversed upon washing the bacteria after three hours and reinstated upon a second induction at six hours (Figure 5.3). Generally, the effect on CFU count was delayed and of lower magnitude than seen with the higher copy vector fig:BB01gc.

The increase in CFU after washing away inducer does not allow us to determine whether some cells were in a bacteriostatic state and recovered, or if the increase in CFU count was simply from bacteria that had yet to die. That is, recovery could be an individual or a population effect. Overall, this demonstrated that expression of barnase could be repressed and depressed by removing IPTG from the medium, and that lower copy number vectors could be used to decrease apparent toxicity of barnase at a population level.

Expression of barnase at the protein level

This expression system was validated at the protein level. Cells were induced in mid-log phase and sampled at 10, 30, and 60 minute intervals for three hours (Figure 5.6). As controls, a full time course of uninduced CSBB01 and induced CSP01 cell extracts were also sampled (Figures C.16 and C.17). Expression of barnase as a protein was tested by separating cell extracts with SDS-PAGE, using a 4-20% gradient gel. Samples were normalized by running equal volumes on a gel, and calibrating protein content using the computer program GeneTools.

A barnase-sized protein increased in relative amount from 20 minutes onward in induced bacteria (arrow) (Figure 5.6). It is similar in size to mature barnase (12.4 kDa, (Hartley and Smeaton 1973)), without the ~2.3 kDA PhoA signal sequence, which is cleaved during membrane translocation (Paddon and Hartley 1987). This suggests that the majority of the barnase is being secreted to the periplasm. Interestingly,

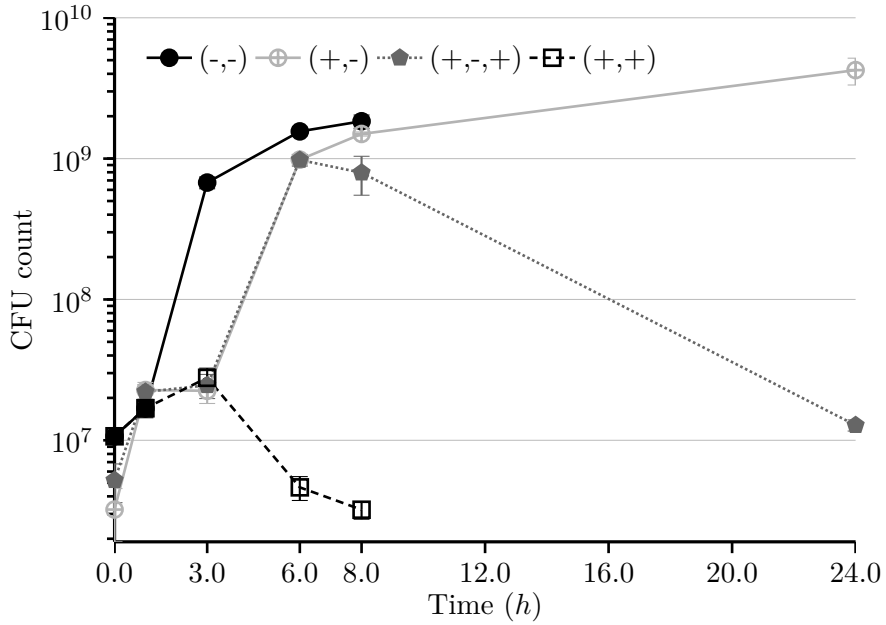


Figure 5.5: Resuscitation of *E. coli* growth after loss of barnase inducer, with barnase on a low copy number plasmid. Cultures of CSBB03 were sampled and plated on antibiotics to determine the number of CFUs at time intervals indicated. At one hour, all but (-●-) (–,–) were treated with 1.0mM IPTG to induce barnase expression. All samples were washed at T3 and resuspended in appropriate medium. (-⊕-) (+,–) Cells resuspended in –IPTG medium. (-⊡-) (+, –, +) Cells resuspended in –IPTG medium, then re-exposed to IPTG at T6. (···⬟···) (+,+) Cells resuspended in +IPTG medium. Measurements are the result of three independent trials, with two replicas per trial, ± SE.

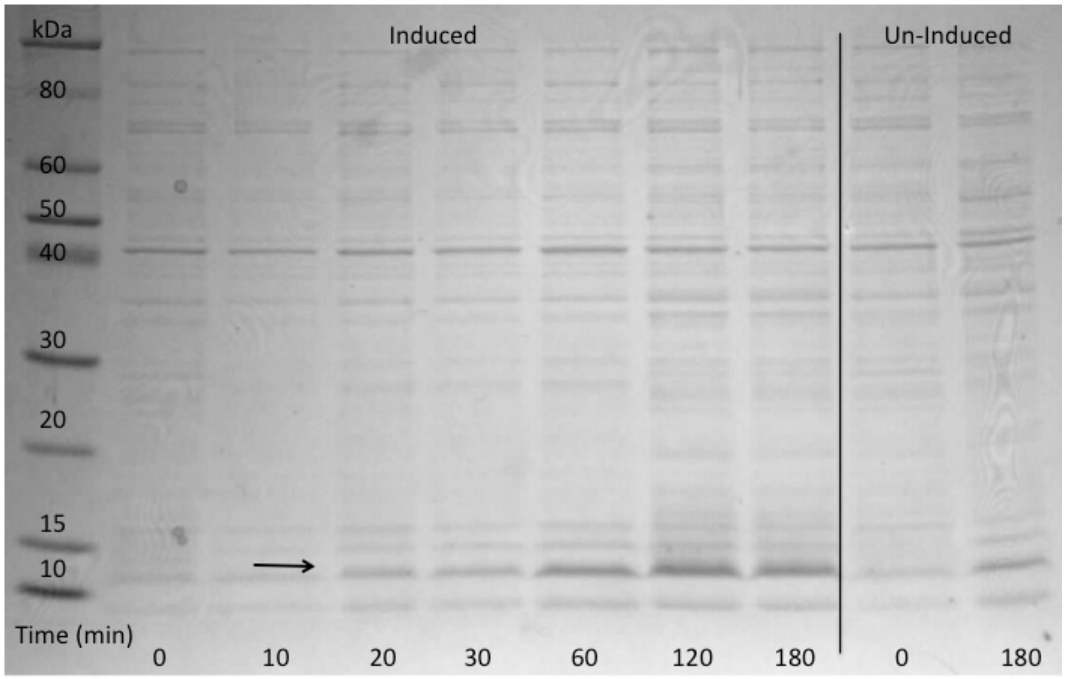


Figure 5.6: Intracellular protein profile of *E. coli* expressing barnase under the P_{tac} promoter. *E. coli* strain CSBB01 was cultured to ~0.500 OD_{600} before sampling, with and without induction (1.0 mM IPTG) of barnase. Protein was extracted from bacteria and separated using SDS-PAGE. Protein content was normalized between samples, and analyzed for presence of barnase (arrow).

barnase also appears to accumulate in uninduced CSBB01 over time, though not to the same degree as in induced bacteria (Figure 5.6). A band for barstar (10.2 kDa) is not apparent.

5.1.2 Expression systems with barnase under the P_{BAD} promoter

To increase our range of possible barnase expression systems, the *phoA*-barnase:barstar region of pMT416 was amplified and inserted into the pBAD-Topo vector, placing barnase under the P_{BAD} promoter (Figure 5.7). This became plasmid pBB02. The promoter is less inducible than P_{tac} but considerably more repressible (Guzman *et al.* 1995).

P_{BAD} is induced by arabinose and repressed by glucose (Guzman *et al.* 1995; Lobell and Schleif 1990). Expression is modulated by the AraC protein, which is both a positive and a negative regulator (Schleif 2010; Lobell and Schleif 1990; Guzman *et al.* 1995). In the absence of arabinose, AraC forms a dimer that binds two non-adjacent sites upstream of the promoter, forming a DNA loop and blocking transcription (Lobell and Schleif 1990). In the presence of arabinose, AraC forms a dimer that binds two adjacent sites just upstream of the promoter, and acts as an activator (Lobell and Schleif 1990; Schleif 2010).

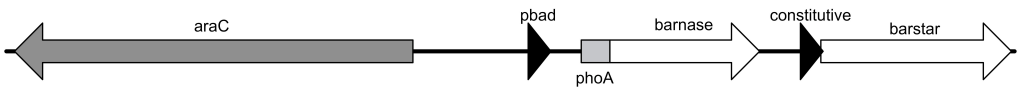


Figure 5.7: Organization of barstar and barnase genes on plasmid pBB02. The RNase barnase is under control of the *BAD* promoter, which can be induced with arabinose at the transcriptional level and repressed by glucose. The signal sequence *phoA* targets the protein for secretion into the periplasmic space in *E. coli*. Downstream, the intracellular inhibitor barstar is under a constitutive promoter. The *araC* gene is a control gene, necessary for the system to be turned on or off. Genes are on a high-copy number pUC-derived backbone.

The *E. coli* strain LMG194, which is *araBAD*–, was transformed with the plasmid and tested for effect on growth as measured by optical density and CFU count. The deletion of the arabinose operon in the chromosome prevents the cell from

metabolizing the inducing substrate (Guzman *et al.* 1995). The plasmid pP02, which expresses a *lacZ* fragment under the P_{BAD} promoter was used as a control. Induction of barnase with 0.2% arabinose under this system resulted in cessation of growth as measured by optical density (Figure 5.8), with growth significantly different than the uninduced control after four hours (p=0.007). This is similar to what was seen with the P_{tac} promoter. Reducing arabinose in 10x increments relieved growth depression of the culture with the barnase construct until it was similar to 'no-arabinose' controls (Figure 5.9). Addition of 0.1 mM arabinose or more resulted in significant growth differences ($p \leqslant 0.001$) at four hours.

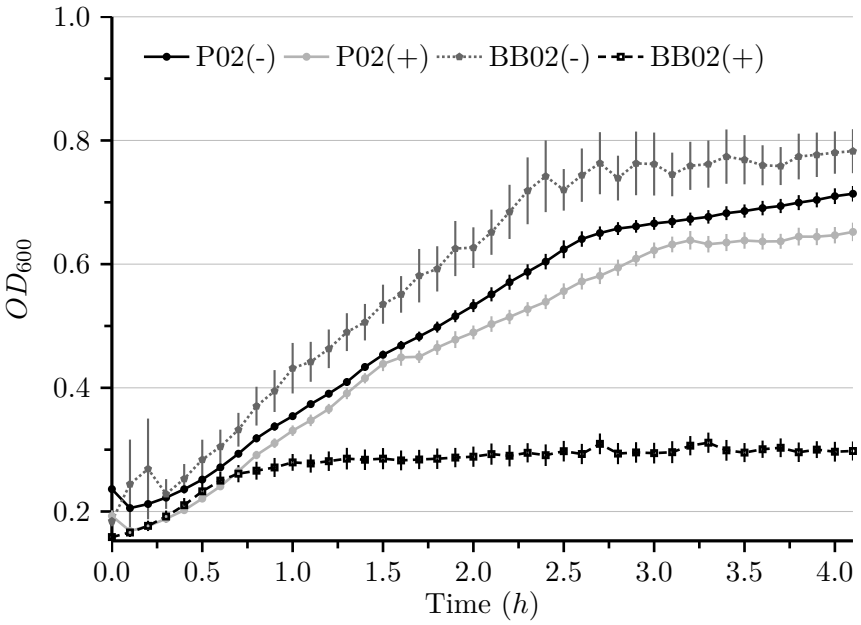


Figure 5.8: Optical density of *E. coli* culture expressing barnase under the P_{BAD} promoter. Cultures were grown to density of ~0.100 OD_{600} before treatment with 0.2% arabinose. Time is given as hours after induction. (-●-) LMP02(–), uninduced . (-⊕-) LMP02(+), *lacZ* induced. (···⬟···) LMBB02(–), uninduced. (-⊟-) LMBB02(+), barnase induced. Measurements were the result of three independent trials with four replicas per trial, ± SE.

Serial dilutions were taken every 30 minutes to determine the affect of arabinose on CFU/ml (Figure 5.10). Unlike barnase expression under the P_{tac} promoter, CFU count did not decrease drastically upon induction, though the two conditions were still significantly different at four hours (p=0.004). Here, the viable cell count remained

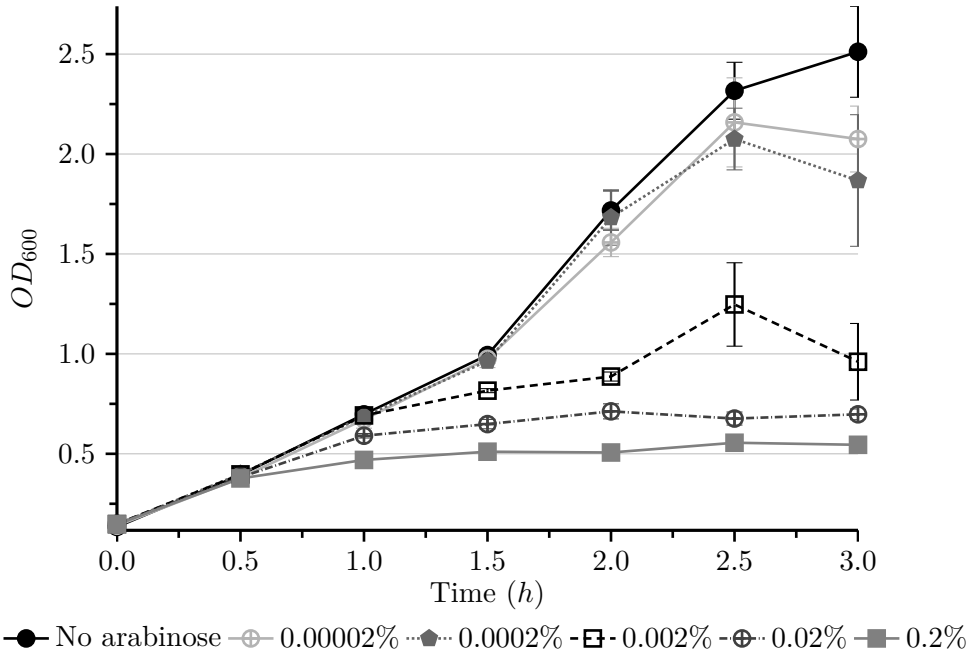


Figure 5.9: Optical density of *E. coli* culture expressing barnase under the P_{BAD} promoter with various concentrations of arabinose. The effect of decreasing arabinose ten fold was tested on the growth of LMBB02 cultures over three hours, as measured by OD_{600}. Time is given as hours after induction. Cells were treated with (—●—) no arabinose. (—⊕—) 0.00002% arabinose. (···⬟···) 0.0002% arabinose. (-⊟-) 0.002% arabinose. (-⊕···) 0.02% arabinose. (—■—) 0.2% arabinose. Measurements were the result of three independent trials with four replicas per trial, ± SE.

steady, with a slight increasing trend, despite being on a high copy number vector. As with the barnase reinduction test, it is impossible to know from this data if this is a cell or population level effect. It is possible that all cells are dividing at a minimal level or that a similar number of cells are dying as are dividing.

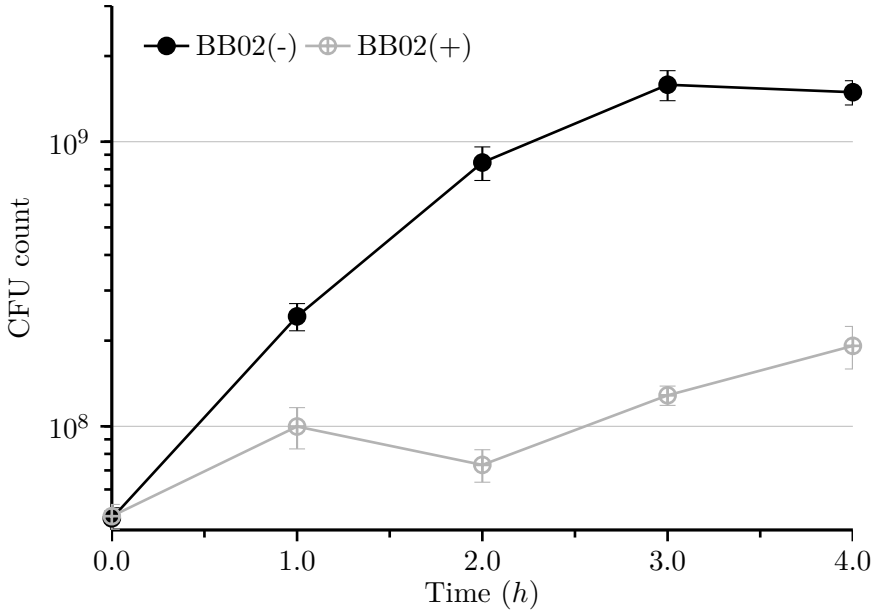


Figure 5.10: Effect of barnase induction under the promoter P_{BAD} on *E. coli* growth. Strain LMBB02 was cultured to 0.1 OD_{600} then split before treatment (time zero) with or without 0.1 mM IPTG. Samples were diluted and plated on selectable medium to enumerate CFUs. (—●—) LMBB02(–), uninduced. (—⊕—) LMBB02(+), induced. Measurements were the result of three trials with two replicas per trial, ± SE.

Barstar protection

To achieve control over barstar expression levels, barstar was synthesized under the P_{tac} promoter, into a high copy number vector that conferred kanamycin resistance, giving us pBS01. Concomitant of barstar had the potential to increase the amount of barnase that could be expressed in the cell without deleterious effects. A two-plasmid system, co-selecting for pBS01 and a barnase-expressing plasmid such as pBB01 or pBB02, did not yield consistently reproducible results. The two plasmids were under the same copy control mechanisms, and, as seen previously, copy number affected

toxicity. The culture may have become quickly populated with bacteria containing a minority of barnase plasmid, just sufficient for ampicillin resistance. To counter any problems due to copy number control, barstar under P_{tac} was inserted into pBB02, putting both the inducible barnase and barstar on the same construct. This gave the plasmid pBB05 (Figure 5.11).

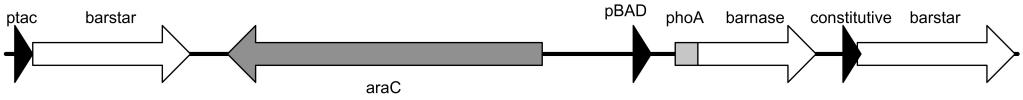


Figure 5.11: Organization of barstar and barnase genes on plasmid pBB05. The RNase barnase is under control of the BAD promoter, which can be induced with arabinose at the transcriptional level and repressed by glucose. The signal sequence *phoA* targets the protein for secretion into the periplasmic space in *E. coli*. The intracellular inhibitor barstar is present in two copies: under the *tac* promoter and under a constitutive promoter. *AraC* is a control gene, capable of up and downregulating expression from the promoter. Genes are on a high-copy number pUC-derived backbone.

The inducible barstar and barnase fragment from pBB05 was successfully cloned into a new plasmid to remove the constitutively expressed barstar, giving pBB06. The resulting strain was either slow growing or did not grow at all in liquid culture, so subsequent work was carried out with pBB05. This is similar to the expression system used in Ramos *et al.* (2005), where both a constitutive and an induced barstar gene were present. The plasmid was tested in strain CSH104, part of the same series of Cold Spring Harbour strains as CSH100. CSH104 is a non-$lacI^Q$ strain, which allows for leaky expression of barstar.

As with pBB02 containing bacteria, induction of barnase from pBB05 led to a decrease in growth of the culture (Figure 5.12) compared to the uninduced strains. Barstar could not protect against barnase if the genes were induced with IPTG together, possibly because the barnase was degrading barstar RNA before a significant pool could be translated. Parallel cultures were grown with or without barstar inducer, IPTG, overnight and re-cultured in the same medium before barnase induction. Cultures just expressing barstar had slightly reduced growth levels when compared

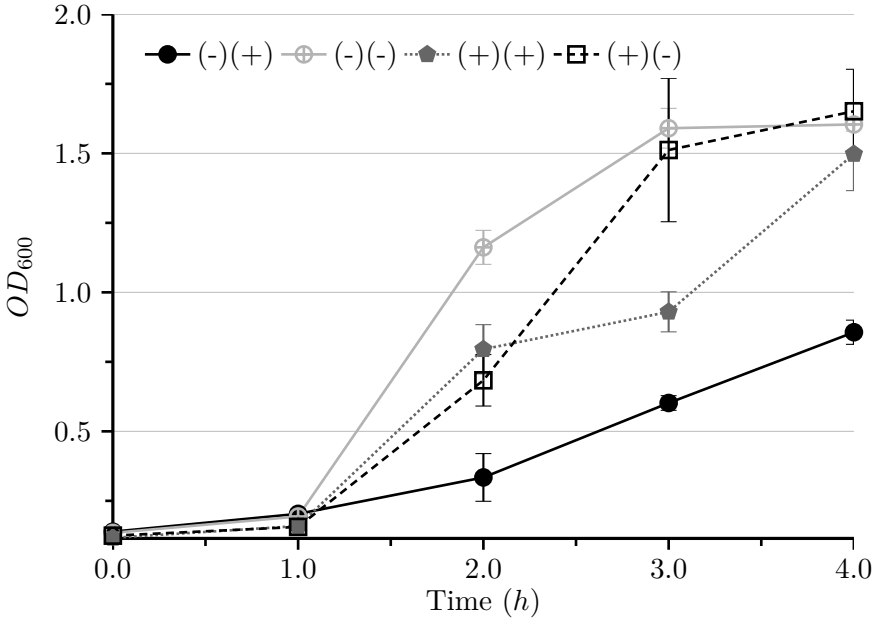


Figure 5.12: Pre-induction of barstar reduces barnase toxicity. Growth was measured as optical density (OD) at 600nm. *E. coli* strain C4BB05 was cultured to ~0.100 OD_{600} before induction of barnase (0.2% arabinose), at T0. (⊕) (–,–) Neither barstar nor barnase induced. (●) (–, +) Just barnase induced. (⊡) (+,–) Barstar pre-induced, by culturing bacteria in barstar-inducing medium overnight (0.1mM IPTG) and during growth experiments. (⬟) (+,+) Barstar pre-induced, followed by barnase induction. Measurements are the result of two independent trials with four replicas within trial, ± SE.

to the controls. This was also seen in cultures transformed with pBS01, an affect more marked in some strains, such as DH5α. Pre-induction of barstar did, however, reduce the loss of growth seen when barnase was induced. This suggests both that barstar was being expressed as a functional protein, and that it would be possible to express barnase (necessary for PSK) without killing the cell.

A final expression system, pBB10, was developed to test barstar and barnase for PSK to have a system where the barnase would remain in the cytoplasm for PSK. The plasmid pBB10 is identical to pBB05, but lacks the *phoA* signal sequence for export of the protein into the periplasm. During development cells were maintained in barnase-repressing conditions, though, as with pBB06, transformants were difficult to work with. Development of these test systems highlighted difficulties in controlling toxicity at levels necessary for creating a PSK-exhibiting environment. Further analysis of the pBB10 system was delayed, and the stability and relative levels of toxin and antitoxin necessary for the PSK phenotype were analyzed.

5.1.3 Relationship between protein expression and degradation necessary for PSK

In the process of optimizing various expression systems for barstar and barnase, the question arises of what the optimal levels of toxin and antitotxin in an individual cell are for generating PSK. Work on TA gene regulation to date has focused on chromosomal loci, when the genes are still present and can be subsequently up or down regulated (Cataudella *et al.* 2013; Cataudella *et al.* 2012). The focus here is on protein dynamics after plasmid, and consequently, toxin and antitoxin gene loss.

Developing models for toxin and antitoxin decay upon plasmid loss

A preliminary formalization of the question was performed, using a short mathematical model. For the model, the degradation of the toxin and antitoxin can be visualized as two intersecting decay functions (Figure 5.13).

$$\mathrm{N_{T0}}2^{-t/\mathrm{H_T}} - \mathrm{N_{A0}}2^{-t/\mathrm{H_A}} > \mathrm{X}$$

A preliminary model for the point of PSK can be seen in the equation below, where t is time, N is starting population of the proteins, H is half-life and X is the number of excess toxin molecules necessary for PSK. The point Y represents the

number of antitoxin molecules in the cell at the time in which PSK is possible. The antitoxin and toxin decay at different rates (Figure 5.13). PSK is possible when the difference between the two populations reaches X. Somewhere within this range are the stoichiometric dynamics that support PSK.

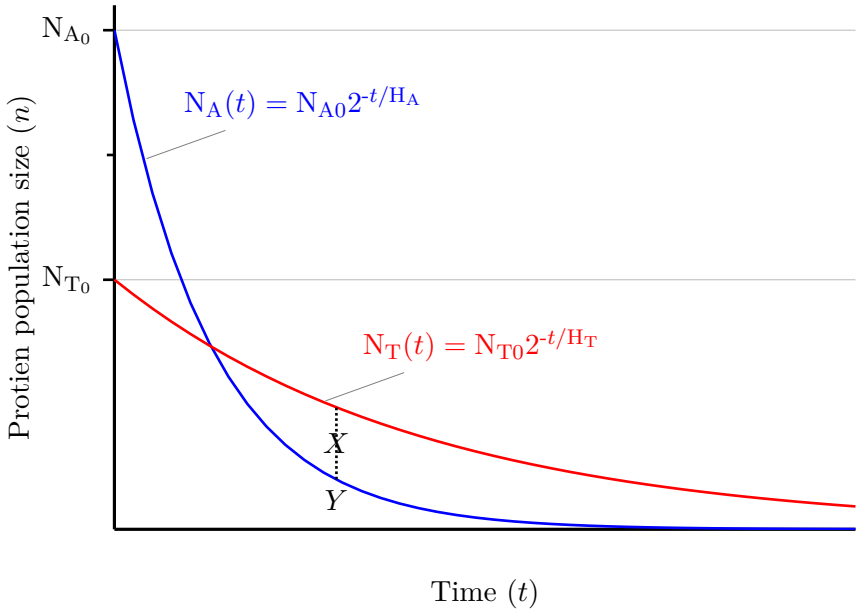


Figure 5.13: Decay of antitoxin and toxin after plasmid loss, with subsequent PSK. Degradation of the toxin (N_{T_0}, red) and antitoxin (N_{A_0}, blue) after plasmid loss (t) is modeled as logarithmic decay, $N2^{(-t/H)}$. PSK is possible when the difference of the two populations exceeds X, the amount of excess toxin necessary for cell death.

Assumptions of this model

1. Logarithmic decay.

   Half-lives are a limited representation of decay within biological systems. It assumes, for example, that the rate of decay of the molecule is the same for the entire lifetime of the protein pool. Yet the stability of type II TA system antitoxins increases upon binding their cognate toxin (Diago-Navarro *et al.* 2013; Melderen *et al.* 1994; Van Melderen *et al.* 1996; Brzozowska and Zielenkiewicz 2013), so that that the rate of decay of the population is likely to change upon depletion of the pool of free antitoxin (after the intercept of the two decay curves). Further, cell division will halve the amount of protein in cells after

the genes are lost, such that the rate of cell division will effect the rate of decay of both the toxin and antitoxin. Cell division will also cause fluctuations in protein concentrations, with some daughters potentially experiencing wide variations in the ratio of toxin to antitoxin.

2. A 1:1 stoichiometry of toxin and antitoxin binding.
   This is true of barstar and barnase, but many type II TA systems bind as multimers (Afif *et al.* 2001; Diago-Navarro *et al.* 2013; Grønlund and Gerdes 1999; Cherny *et al.* 2005; Dao-Thi *et al.* 2000; Cherny *et al.* 2007). For these cases, each multimer of antitoxin, necessary to inhibit the activity of a functional multimer of toxin, can be considered a 'unit'. N becomes the number of functional units of antitoxin and toxin within the cell, such that PSK is possible when the number of toxin units exceeds the specified number of antitoxin units.

3. $N_A > N_T$ at t=0.
   At the point of plasmid loss, there must be more antitoxin in the system than toxin, to avoid plasmid suicide. Again, N must be considered a functional unit, as this will not be strictly true if an antitoxin molecule binds multiple toxins.

4. $H_A < H_T$.
   For PSK to occur, the antitoxin units must decay faster than toxin units. Because the half-life of some toxins is primarily a function of cell division while antitoxin decay is primarily proteolytic, a change in cell division will change the relative rate of toxin to antitoxin decay.

5. Toxicity at X .
   There is some number of toxin molecules, X , in excess of the antitoxin necessary to affect cellular function. This means that the antitoxin population must drop below the toxin population by a minimum amount.

6. Time.
   While the antitoxin levels may drop below the level of toxin given an infinite time period, it is assumed that there is a fixed time period within which PSK can be effective. Beyond this time, the plasmidless cell may revive.

Applying equations of logarithmic decay to determine relative expression of barstar over barnase for optimal PSK.

This equation provides a rough estimate of the expression levels necessary for PSK with a toxin and an antitoxin of a given stability, but relies on *in vivo* data. The differential lability of barstar and barnase in the cellular environment is unknown, though their stability *in vitro* has been well characterized.

For the sake of illustration, I use available *in vitro* data with a number of assumptions, in particular, that their relative stability *in vivo* was proportional to their relative stability *in vitro*. Their free energy of unfolding (ΔG_u), is 5.2 ± 0.05 kcal/mol and 8.95 ± 0.05 kcal/mol, as determined by denaturation in urea (Khurana and Udgaonkar 1994; Pace *et al.* 1992; Serrano *et al.* 1992).

If barnase were metabolically stable enough so that its half-life was effectively determined by cell division (⩾30 minutes), then barstar would have a half-life of 17.52 minutes based on their relative ΔG_u. If the number of molecules necessary for cell death is set at 1, for PSK to occur in less than five cell divisions (t<150 minutes), the equation becomes:

$$\begin{aligned} \mathrm{N_{T0}}2^{-t/\mathrm{H_T}} - \mathrm{N_{A0}}2^{-t/\mathrm{H_A}} &< \mathrm{X} \\ (100)2^{(-150/30)} - \mathrm{N_{A0}}2^{(-150/17.52)} &> 1 \\ \mathrm{N_{A0}} &< 863 \end{aligned}$$

This gives us a ratio of 8.6 antitoxins to 1 toxin molecule in the cell, to get PSK at 150 minutes. The lower limit is 1:1 (discussed above), meaning that there would have to be between 1 and 8.6 times the amount of antitoxin in the cell at the time of plasmid loss to get PSK. To get an effect within two cell divisions (t<60 minutes), the amount of antitoxin would have to be 2.57 times the toxin. The ratio of expression would have to be higher, due to antitoxin degradation during cellular growth.

This is primarily an illustration of the parameter space of the equation: the link between *in vivo* and *in vitro* stability is unlikely to be proportional. Degradation of type II TA system antitoxins is protease mediated (Brzozowska and Zielenkiewicz 2013), which target unstructured regions not seen in barstar (Lubienski *et al.* 1994). If barstar proves to relatively more stable *in vivo*, how does this affect the necessary

expression levels? In other words, to get PSK within a defined time point, how does N_A at plasmid loss (N_{A_0}) change when H_A moves closer to H_T?

Relationship between protein abundance and protein stability for generating PSKs

The relationship between antitoxin stability and expression is diagrammed below, with functions for the decaying toxin and the antitoxin, and the difference between the two (Figure 5.14). Here, X is represented as a dashed line: PSK is possible when the the difference between the number of toxin and antitoxin (brown line) crosses X.

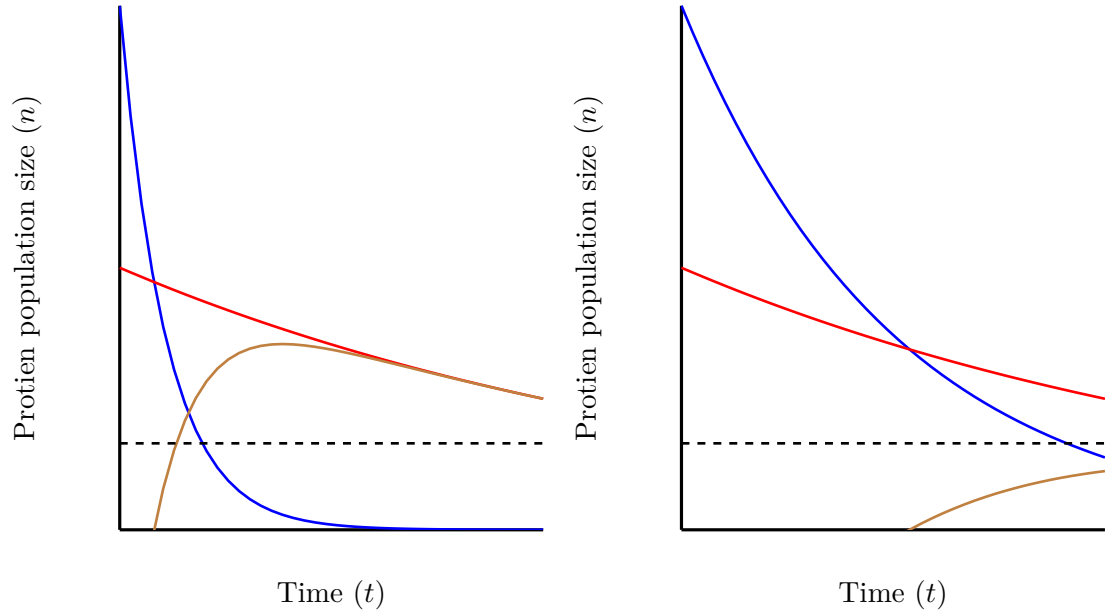
High levels of antitoxin protect the cell when plasmid is retained, and rapid decay after plasmid loss increases the chance of PSK (Figure 5.14a). The difference (brown line) quickly crosses the X. If the antitoxin were more stable, a large pool of slowly decaying antitoxin would delay PSK after plasmid loss, and the toxicity point may not be reached within a given time period (or not at all) (Figure 5.14b). Low relative expression of a highly labile antitoxin means the number of antitoxins could easily drop below the number of toxins when the plasmid is retained, increasing the chance of plasmid suicide (Figure 5.14c). Finally, though a stable antitoxin expressed at low levels means that toxicity may be reached and PSK will occur (Figure 5.14d) it is more marginal than seen with a large population of unstable antitoxin.

These differences can be summarized by analyzing the effect of changing one parameter, antitoxin stability, on the rest of the system. The remaining terms (X, t at PSK, H_T, N_{T0}) are fixed. Assuming a set toxin decay curve and a set time to PSK, there will be a fixed coordinate (Y, T) that the antitoxin decay curve must pass through (Figure 5.13).

As H_A approaches H_T, N_{A_0} approaches N_{T_0} Under conditions where PSK could occur.

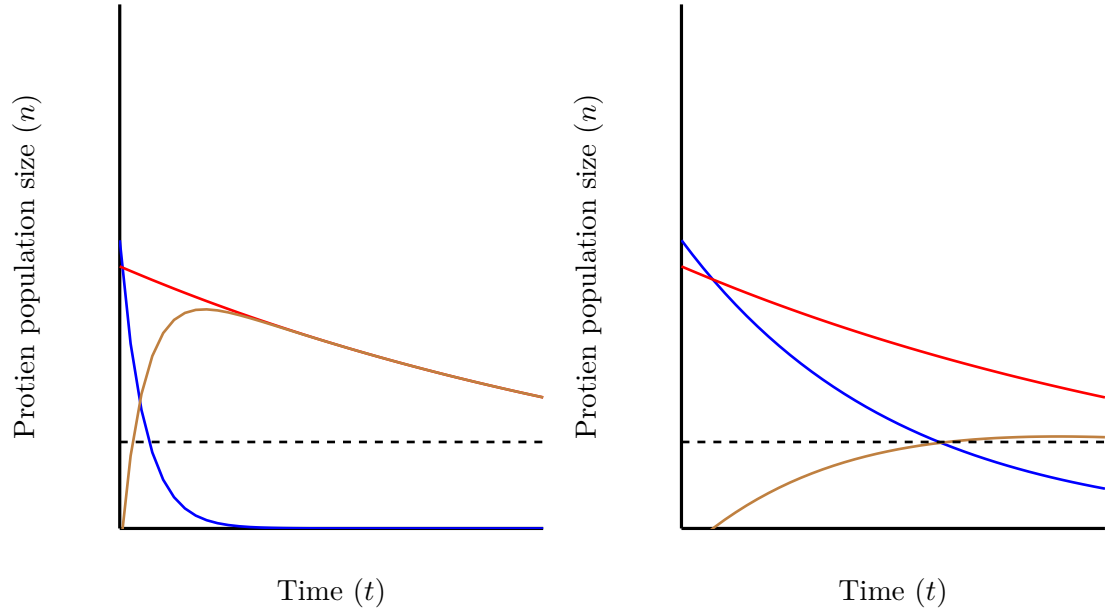
Refering to (Figure 5.13), with the population of antitoxin ($N_A(t)$) and toxin ($N_T(t)$) in the cell as:

$$N_A(t) = N_{A_0} 2^{-t/H_A}$$
$$N_T(t) = N_{T_0} 2^{-t/H_T}$$



(a) Large population of unstable antitoxin.

(b) Large population of stable antitoxin.



(c) Small population of unstable antitoxin.

(d) Small population of stable antitoxin.

Figure 5.14: Changing the relative starting populations and rate of decay of antitoxin in respect to a fixed toxin. Decay curves of the antitoxin (blue), and toxin (red) presented, with the difference between the two (brown) such that PSK occurs when it crosses X (black). Starting population of antitoxin is either double (large starting population) the toxin or just above (small population). The antitoxin is either degraded ~3 times as fast as the toxin (unstable) or twice as fast (stable).

As stated previously, the conditions necessary for PSK impose the following assumptions:

1. The antitoxin in the cell at $t = 0$ is greater than the toxin:

$$\mathrm{N_{A_0}} \geqslant \mathrm{N_{T_0}}$$

2. The half life of antitoxin is less than the half life of the toxin.

$$\mathrm{H_A} \leqslant \mathrm{H_T}$$

We choose a fixed time, $t = T$ and an X such that

$$\mathrm{N_T}(T) - \mathrm{N_A}(T) = X > 0$$

which exists by our assumptions.
Without loss of generality we fix $\mathrm{N_T}(t)$ and Y defined as:

$$Y = \mathrm{N_A}(T) = \mathrm{N_{A_0}} 2^{-T/\mathrm{H_A}}$$

Which is rearranged to:

$$\begin{aligned} \mathrm{N_{A_0}} &= \frac{Y}{2^{-T/\mathrm{H_A}}} \\ &= Y 2^{T/\mathrm{H_A}} \end{aligned}$$

We define $\widetilde{\mathrm{H}}_\mathrm{A}$, closer to $\mathrm{H_T}$, as:

$$\widetilde{\mathrm{H}}_\mathrm{A} = \mathrm{H_A} + \epsilon$$

where

$$0 < \epsilon < \mathrm{H_T} - \mathrm{H_A}$$

and define

$$\widetilde{\mathrm{N}}_{\mathrm{A}_0} = Y 2^{T/(\mathrm{H_A}+\epsilon)}$$

Now

$$\begin{aligned} \delta_{\mathrm{N_{A_0}}} &= \mathrm{N_{A_0}} - \widetilde{\mathrm{N}}_{\mathrm{A}_0} \\ &= Y \left(2^{T/\mathrm{H_A}} - 2^{T/(\mathrm{H_A}+\epsilon)}\right) \end{aligned}$$

and as

$$\frac{T}{\mathrm{H_A}} > \frac{T}{\mathrm{H_A} + \epsilon}$$

then

$$\delta_{\mathrm{N_{A_0}}} > 0$$

Conclusion: For toxicity within a defined time point, as H_A approaches H_T, N_{A0} approaches N_{T0}.

From this, as the antitoxin becomes more stable, then the starting population of antitoxin must approach the level of toxin to get enough free toxin molecules to allow for PSK within a given time point. This analysis primarily focuses on the relationship between toxin and antitoxin protein populations and their decay. But similar equations can be used to explore the effect of changing t and X. It is evident that to reduce t, the time between plasmid loss and PSK, either $\mathrm{H_A}$ needs to decrease or $\mathrm{N_A}$ needs to approach $\mathrm{N_T}$. If the toxicity of the toxin were to decrease (X increased), the antitoxin would also need to decay faster or have a smaller starting population to allow for PSK within the same time period.

Many of the observations here are based on the relative stability and population of the two proteins. Often, increasing $\mathrm{N_{A0}}$ would have a similar effect to decreasing $\mathrm{N_{T0}}$. But given assumptions 5 and 6 (above), there needs to be a specific number of toxin molecules free within a given time period: double the amount of toxin over antitoxin would not be sufficient if there is only one antitoxin molecule and three toxin are necessary for affecting the cell. Relative numbers also do not take into account metabolic limitations. Actual $\mathrm{N_{A0}}$ and $\mathrm{N_{T0}}$ will be influenced by the costs of protein production on the cell.

5.2 Discussion

I am interested in predicting when toxin-antitoxin gene pairs can create the conditions necessary for PSK. In particular, what differentiates a toxin and antitoxin gene pair that creates a PSK phenotype from those that do not. The toxin and antitoxin gene pair barnase and barstar from *B. amyloliquefaciens* served as a model system. They are not known to exhibit PSK. In this chapter I explored what conditions were necessary for them to do so.

5.2.1 Various expression systems compatible with barstar and barnase, but may not be sufficient for a PSK effect

Toward this aim, I constructed a range of expression systems with which to test barstar and barnase as a PSK in *E. coli*. The proteins were induced under P_{tac} and P_{BAD}, in high and mid-copy number vectors. Under all conditions, barnase expression resulted in reduction of culture growth. The highest reduction in growth was seen with barnase expressed from P_{tac}, in a high copy number vector. Use of the P_{BAD} promoter and lower copy vectors reduced this effect. Studies of alkaline phosphatase expression under P_{BAD} and P_{tac} suggest that P_{tac} had a 12-40x smaller ratio of protein induction to protein repression, primarily due to less repression (Guzman *et al.* 1995). Expression systems utilizing a *lac*-derived promoter on plasmids are notoriously leaky (Siegele and Hu 1997). The reduction in growth was reversible, with cultures transformed with both vectors resuming growth when IPTG was removed from the media.

As seen in all growth curves, expression of barnase limited cell growth. Yet some expression of the toxin during cell growth is necessary to get PSK after plasmid loss. A system was developed where barstar was inducible under the P_{tac} promoter, to increase levels of antitoxin in the cell and thus increase tolerance of the cell to barnase. With this system, both barstar and barnase could be induced, giving higher levels of growth than when barnase was induced alone. Such two promoter systems can be complicated by the fact that IPTG, necessary for derepression of P_{tac}, interferes with induction of genes under P_{BAD} (Lee *et al.* 2007). This effect was ameliorated by using high levels of arabinose to induce barnase expression (Lee *et al.* 2007).

Different concentrations of arabinose gave a different effect on cell growth, with 0.00002% arabinose causing almost no growth cessation. Guzman (2005) found a 300-fold difference in reporter protein (alkaline phosphatase) production using different concentrations of inducer. Yet the effects of protein expression as measured here were population measurements, not individual cell measurements. It is impossible to ascertain with these methods if *each* cell is expressing intermediate concentrations of protein (cell level effect), or if differing proportions of the bacteria are fully induced while the rest are not (population level effect).

The expression of P_{tac} and P_{BAD} have been investigated at the cell level, using GFP expression as a measure of induction. Siegele *et al.* found that bacteria ex-

pressing GFP from P_{BAD} had a bimodal distribution, suggesting that it is an on or off mechanism at the cell level. This is similar to what is seen with the lac operon (Siegele and Hu 1997; Novick and Weiner 1957). In the case of the dual promoter system, pBB05, it may be that there are four subpopulations: one with no induction, a second and third with either barstar or barnase induced, and a fourth with both. For both promoter systems, high levels of inducer (as used here) increase the proportion of the population in the 'on' mode (Lee *et al.* 2007; Siegele and Hu 1997).

Though not used here, it is also possible to get intermediate expression from P_{tac} by using strains that are deleted for permeases (Siegele and Hu 1997; Khlebnikov *et al.* 2000). Random spontaneous induction of the operons in the absence of inducer creates a population of cells with a range of permease/uptake protein levels. At low levels of inducer, only some cells will contain enough permease to bring in the inducer and turn on gene expression for more permease, creating the bimodal population structure (Siegele and Hu 1997; Khlebnikov *et al.* 2000). Strains deleted for permeases result in a more homogenous population with intermediate expression levels, though they are still bimodal at the cellular level.

Future work for engineering PSKs would need to ensure that protein expression was being controlled on the individual cell level- it is likely that the best mechanisms would involve within cell regulation as opposed to reliance on external inducers. Experimental systems used will also require knowledge of the intracellular stability of the toxin and antitoxin, to determine the desired expression target.

5.2.2 Level of protein expression dependent on respective stability of toxin and antitoxin, which may affect PSK evolution

This work focuses on conditions necessary for barstar and cytoplasmic barnase to successfully mediate PSK in a manner similar to type II TA systems. Effective PSK by type II TA systems relies on two conditions: that there be enough antitoxin in the cell for it to survive while the plasmid is retained, and that the antitoxin is lost faster than the toxin after plasmid loss. Determining necessary amounts of barstar and barnase expression relies on the speed at which the proteins are degraded.

Using equations of logarithmic decay, it is evident that as the stability of the antitoxin increases, the population of antitoxin at plasmid loss must get closer to the level of toxin to get PSK within a defined time point. From this, a general

schematic can be derived on the relationship between the relative stability of the toxin and antitoxin, and their relative expression in the cell prior to plasmid loss (Figure 5.15).

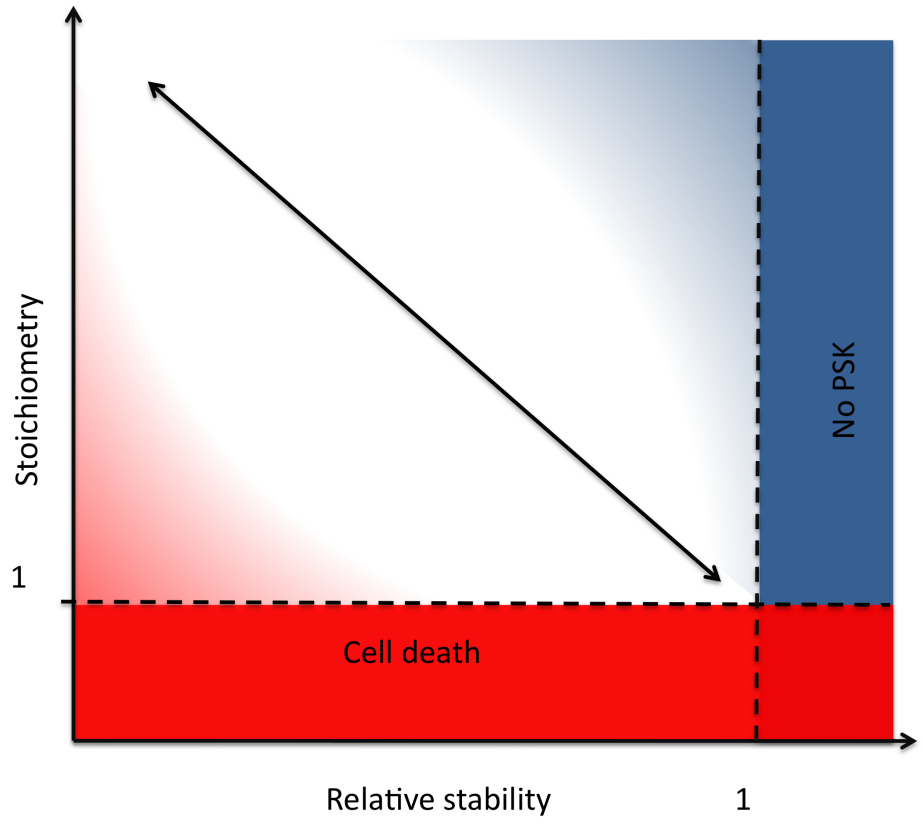


Figure 5.15: Relationship between the relative stability and stoichiometry of toxin and antitoxin necessary for PSK generation. The ratio of antitoxin to toxin (A:T) (vertical axis) must be greater than one when the plasmid is retained to avoid plasmid suicide (red). The chance of this increases as the ratio nears one, due to noise in biological systems. The relative stability of the antitoxin to toxin (horizontal axis) must be less than one, otherwise there will be no PSK phenotype (blue). The chance that there will be enough free toxin to cause PSK after plasmid loss decreases as the number approaches one. The amount of antitoxin in excess of toxin necessary for PSK varies with stability (arrow). As stability of the antitoxin increases, the ratio of A:T decreases to retain PSK.

On the vertical axis, if ratio of antitoxin to toxin is below one, there will be cell death before plasmid loss (plasmid suicide). The chance of this is likely to increase as the ratio approaches one, due to considerable stochasticity within biological systems. Gene expression systems are inherently noisy, due to within cell (intrinsic) factors such as transcription and translation and between cell (extrinsic) factors such as cell cycle (Raser and O'Shea 2005; Rao *et al.* 2002). Relying on small numbers of excess antitoxin, particularly highly unstable antitoxin, is more likely to put the plasmid at risk of suicide.

On the horizontal axis, if the antitoxin is more stable than the toxin, there will not be PSK after plasmid loss. The chance of this increases as the ratio of antitoxin to toxin stability approaches one, as stable antitoxins are less likely to degrade swiftly enough to get PSK within a physiologically relevant time frame. This is particularly true of stable antitoxins expressed at high levels.

The arrow indicates the conditions of successful PSK, where the starting ratio of the antitoxin to toxin decreases as the antitoxin becomes more stable. Conditions for PSK can occur at high relative stability and a low initial ratio of antitoxin to toxin in cell, as illustrated in the bottom right of the graph. But it requires conditions that make it sensitive to both too much antitoxin (no PSK) or too little (premature cell death). TA systems that tended toward this corner would be more sensitive to small changes in relative toxin and antitoxin expression during cellular growth.

This is likely to be an issue for TA systems on mobile elements, as they move into new hosts. Noise is known to affect expression within cells within clonal populations. PSKs on mobile elements, particularly those with a broad host range would be subject to significantly larger fluctuations in gene expression due to strain and species-dependent differences. I speculate here that there would be an evolutionary trend toward PSK systems that operate under the conditions described in the bottom left of (Figure 5.15), with high expression levels of unstable antitoxins. This allows for robustness in the face of a noisy and changing cellular environment.

This is seen in known PSK systems (Van Melderen *et al.* 1996; Franch *et al.* 1997). Cells treated with the bacteriostatic antibiotic rifampicin lost their MazE in 30 minutes, while the MazF toxin remained stable over four hours (Aizenman *et al.* 1996). The antitoxin RelE has a half-life about 10-fold less than RelB (assuming a 30 minute cell division), and is present in the growing cell 10-fold higher (Cataudella *et al.* 2013), suggesting it is transcribed 100-fold more than RelB. Though this analysis

is based on type II TA systems, analogous processes would be occurring in type I TA systems. Hok mRNA is 40-fold more stable than the antitoxin Sok RNA (30 seconds, (Franch *et al.* 1997)) and SprA1 mRNA is 20-fold more stable than $SprA1_{AS}$ RNA (10 minutes, (Sayed *et al.* 2012)). $SprA1_{AS}$ is also expressed in 30-80-fold excess of the toxin (Sayed *et al.* 2012). Increasing levels of unstable antitoxin will be limited by metabolic costs to the cell.

Another interesting side observation of this model is the effect that changes in the rate of cell division will have on the effectiveness of PSK. For metabolically stable proteins, like many TA toxins, their effective half-life is equal to the rate of cell division (Cataudella *et al.* 2013). Slower cell division will increase the stability of the toxin and thus the relative decay of the antitoxin. This would be particularly relevant for the response of TA systems to cellular stress, which both slows cell division and increases protease production, further skewing the system toward free toxin.

This is a considerably simplified model. There are number of assumptions associated with logarithmic decay, and the model here is more appropriate for large populations of molecules (Leighton Turner, personal communication). Small populations are subject to noise, which is not accounted for in our model, and the decay of the proteins is likely to be less deterministic and more stochastic (Raser and O'Shea 2005; Rao *et al.* 2002). Dealing with the ratio of the two molecules, rather than the absolute numbers, may be a better system for modeling small populations of proteins as would be expected here (Leighton Turner, personal communication). But, as discussed in the results, the absolute numbers of proteins is important given t and X : enough molecules of toxin must be free within a non-infinite period of time to get PSK.

Applying the model to barstar and barnase

Attempts to determine relevant levels of barstar and barnase in the cell to derive PSK are infeasible at this time. Only the *in vitro* stabilities of barstar and barnase are known. Barstar is less stable than barnase thermodynamically and in urea. Many type II antitoxins are less stable *in vitro* as well as *in vivo*. The antitoxins YefM and PhD, with a T_m below 37℃, are primarily unfolded as translated, taking on a folded structure when bound to the toxin (Cherny *et al.* 2005; Gazit and Sauer 1999). Their respective toxins have a T_m of 60℃ and 65℃(Cherny *et al.* 2005), as

measured by circular dichromism at physiological pH. The antitoxin CcdA has a T_m of 54.2℃ at physiological conditions, but this is still less than the cognate toxin (Dao-Thi *et al.* 2000), which has a free energy of unfolding twice that of the antitoxin. Correspondingly, the CcdA has a half-life of 30 minutes *in vivo* and the toxin CcdB has a half-life of 60 minute(Van Melderen *et al.* 1996) after cell division has been stalled with rifampicin.

This correlation, though, is not always consistent. The type II antitoxin RelB has a T_m of 58.5℃, slightly higher than the cognate toxin's T_m of 52.5℃ (Cherny *et al.* 2007). Yet RelB has a half life of three minutes in the presence of Lon protease *in vitro*, while RelE remained stable (Overgaard *et al.* 2008). Antitoxin stability *in vivo* is typically mediated by proteases Lon and ClpP (Brzozowska and Zielenkiewicz 2013; Diago-Navarro *et al.* 2013; Dao-Thi *et al.* 2000; Van Melderen *et al.* 1996), which target unstructured regions of the protein that become ordered or shielded upon toxin binding (Brzozowska and Zielenkiewicz 2013). These unstructured regions are likely the cause of antitoxin instability *in vitro* as well.

Barstar is relatively less stable than barnase, but has a higher T_m than many type II toxins and is overall a stable protein. Most importantly, though, barstar notably lacks the unstructured regions targeted by proteases (Lubienski *et al.* 1994), making it likely to be even more stable *in vivo* (Dr. Vic Arcus, personal communication). This could be confirmed by tracking fluorescence in fusion proteins in real time or by quantitative Western blots. As discussed, stable antitoxins are difficult to make into PSKs. Over-expression can lead to no PSK and under-expression to premature cell death. Barstar and barnase, despite biochemical similarities to toxins and antitoxins found in TA systems, would at the least require very tight titration of gene expression. This is difficult to achieve using external inducers and modern cloning techniques. This, and the residual toxicity of barstar as well as barnase, suggests that the genes are not ideally suited for engineering into a PSK.

Native TA systems control expression of the toxin and antitoxin by tight autoregulation of the operon during growth. The antitoxin typically has a DNA binding domain that binds and represses the promoter in complex with the toxin (Marianovsky *et al.* 2001; Zhang *et al.* 2003; Kedzierska *et al.* 2007). The binding of the toxin and antitoxin is sensitive to changes in the ratio of the two proteins in the cell, affecting their affinity for the promoter (Cataudella *et al.* 2012; Cataudella *et al.* 2013). This sort of conditional cooperativity makes the system sensitive to changes

in the cellular environment.

It may still be possible to engineer such systems with internally regulated expression systems, tight control of gene expression, and unstable antitoxins. Tripathi *et al.* (2012) was able to increase the relative amount of toxin in the cell by expressing an inactive toxin analogue and titrating antitoxin out of the cell. It is possible though, that these systems would remain sensitive to environmental fluctuations without the tight autoregulation seen in native PSKs (Cataudella *et al.* 2012; Cataudella *et al.* 2013; Nakayama and Kobayashi 1998; Tao *et al.* 1991). Future research into the levels of gene expression and protein stability necessary for PSK may be better done by manipulating TA systems known to exhibit PSK, possibly through mutagenesis of regulatory and protein coding regions.

Chapter 6

Exhibition of PSK by secreted toxin barnase

Most PSK systems described to date consist of a cytoplasmic toxin and antitoxin which are both produced by genes in the bacterium. However, it has been noted that antibiotic resistance genes act similarly to PSK determinants, though the corresponding toxin (the antibiotic) is provided by an exogenous source (Cooper and Heinemann 2000). Consistent with that speculation, PSK has recently been shown for secreted bacteriocins (Inglis *et al.* 2013). In this chapter, I explore the role of genetic context by testing the hypothesis that a secreted toxin can induce a PSK-like activity when expressed from a plasmid that also produces an intracellular antitoxin.

In *E. coli*, secreted barnase has been reported to limit the growth of neighbouring bacteria (Ramos *et al.* 2006); barnase- and barstar-containing plasmids have been described as relatively stable (Ul'yanova *et al.* 2007). These two observations are consistent with some PSK activities. Thus I chose to use barstar and barnase, the latter of which is secreted both in its original host, *B. amyloliquefaciens*, as well as in synthetic constructs used in *E. coli*, to test my hypothesis.

Firstly I looked for indications that a barstar/barnase homologue has been attached to a horizontally mobile element, suggesting a role in evolution by PSK. The potential for the genes to be mobilized was investigated by searching databases of plasmid and virus genome sequences for barnase, barstar and related proteins. Then, I tested barstar and secreted barnase as PSK systems in *E. coli*, including secretion of the toxin, and effect on the growth of neighboring bacteria. The genes were tested for their ability to stabilize plasmids in monoculture, a useful diagnostic for PSK.

6.1 Results

6.1.1 Searches for barstar and barnase on mobile elements

The distribution of barstar and barnase homologues on chromosomes has been analysed before (Condon and Putzer 2002)- I am interested in their presence on mobile elements. The search was widened to include extracellular RNases known to interact with barstar and barnase.

Compilation of interacting RNases

Small, low molecular weight guanyl-specific RNAses similar to barnase have been found in *Bacillus intermedius*, *Bacillus pumilus*, *Bacillus thuringiensis*, *Bacillus circulans*, and *Bacillus coagulans* (Znamenskaya *et al.* 2002; V Znamenskaya *et al.* 1999; Yakovlev *et al.* 1995; Ul'yanova *et al.* 2007). They have 85-100% amino acid sequence identity (Znamenskaya *et al.* 2002), and many can be inhibited by barstar (Yakovlev *et al.* 1995; Ul'yanova *et al.* 2007) (Table 6.1).

The *Streptomyces* ribonucleases Sa, Sa2 and Sa3, have about 22-27% sequence similarity to barnase (Krajcikova and Hartley 2004). Despite low sequence similarity there is high structural similarity in their active sites, and they can form tight associations with, and be inhibited by, barstar (Hartley *et al.* 1996; Ul'yanova *et al.* 2007). Sa, produced by *S. aureofaciens* strain BMK, forms almost identical contacts with barstar as does barnase (Krajcikova and Hartley 2004; Krajcikova *et al.* 1998). The *Sachharopolyspora erythraea* RNase St is not inhibited to the degree of the other RNases, with barstar providing only low levels of protection against toxicity in *E. coli*.

Search on mobile elements

The presence of these toxins and inhibitors on plasmids and viruses was investigated using the amino acid-based search algorithm blastp of the NCBI genomic database. Regions of sequence similarity on different genomes (matches) were filtered to identify putative homologues in virus and plasmid genomes. Narrowing the results by 'virus' did not return any matches. Narrowing the results by plasmid gave only eleven (Table 6.2). Few novel matches were returned from queries of the related RNases

and inhibitors that were not returned for barstar and barnase: those are noted in the table (Table 6.2). The matches were sorted by E-value.

Table 6.1: Microbial RNases and their inhibitors known to interact with barstar and barnase

Species	Protein	Notes
B. amyloliquefaciens	Barnase	Not responsive to phosphate and nitrogen limitation (Znamenskaya *et al.* 1995), though expression may be SpoA dependent (Ulyanova *et al.* 2011). Expressed as a pre-pro-protein, and processed into a mature form by extracellular serine proteases (Paddon and Hartley 1987).
	Barstar	Inhibits barnase intracellularly. Forms a tight complex with barnase (dissociation constant, K_d of $6x10^{-14}$ M (Hartley 1993))
B. Intermedius	Binase	Structural analogue to barnase (Sharipova *et al.* 2005). Inhibited by barstar (Ul'yanova *et al.* 2007; Yakovlev *et al.* 1995), with a $K_d = 4.13x10^{-12}$, 10-fold higher than seen with barstar and barnase under the same conditions (Yakovlev *et al.* 1995). Regulated by phoshate and nitrogen limitation (Sharipova *et al.* 2005; Kharitonova and Vershinina 2009), as well as SpoA (Ul'yanova *et al.* 2007)
	Unknown	An intracellular inhibitor for binase in *B. intermedius* has not been found. It may be extremely unstable (Yakovlev *et al.* 1995), or secretion of binase may occur in such a way to prevent intracellular toxicity (Ulyanova *et al.* 2011).
B. pumilus	Bpu RNase	Inhibited by barstar, reducing toxicity of the RNase in bacterial expression systems (Ul'yanova *et al.* 2007; V Znamenskaya *et al.* 1999). Regulated by SpoA and secreted under phosphate starvation in *B. subtilis* (Ul'yanova *et al.* 2007).
B. circulans	Bci RNase	Secreted under conditions of phosphate and nitrogen starvation (Kharitonova and Vershinina 2009).
Streptomyces aureofaciens	Sa RNase	Inhibited by barstar (Ul'yanova *et al.* 2007; Yakovlev *et al.* 1995), with a K_d of three versions Sa, Sa2 and Sa3 $2x10^{-14}$, $4x10^{-10}$, and $2x10^{-12}$ M, respectively (Hartley *et al.* 1996).
	SaI	Only 29.2% overall amino acid similarity to barstar, but 76.9% similar over the amino acids that bind the respective RNase (Krajcikova *et al.* 1998).
Sachharopolyspora erythraea	St Rnase	Somewhat inhibited by barstar, which provided some protection but not enough for large-scale expression of the RNase in *E. coli* (Hartley *et al.* 1996).

Table 6.1: Microbial RNases and their inhibitors known to interact with barstar and barnase continued

Species	Protein	Notes
	Sti	Barstar homologue, binds barnase with a K_d of $7x10^{-7}$ M. Binds RNase Sa3 with a K_d of $5x10^{-12}$ M.

Table 6.2: Presence of extracellular RNases and inhibitors on plasmids

Species	E- value[1]	Location	Notes
Rnases			
Bacillus cereus YP_001967275	5.56×10^{-34}	pPER272: 272 kb plasmid, periodontal isolate.	Ribonuclease domain of large 1131 aa, putative S-layer protein. Barstar-like protein downstream.
B. Cereus YP_001966638	1.6×10^{-34}	pCER270: 270 kb plasmid, emetic toxin producing strain.	Ribonuclease domain of large 1131 aa, putative S-layer protein. Barstar-like protein downstream.
Burkholderia phymatum YP_001862007	1.36×10^{-22}	pBPHY01	Annotated as a guanine-specific ribonuclease similar to N1 and T1. Found by similarity to Sa, Sa3 and St
Xylanimonas cellulosilytica YP_003327973	9.01×10^{-1}	pXCEL01	No barstar-like protein nearby. Predicted SAF family protein upstream (SM00858 SMART accession number). Found by similarity with binase and bpu
Inhibitors			
B. cereus YP_001967276	1.25×10^{-15}	pCER270: 270 kb plasmid, emetic toxin producing strain.	Downstream of a barnase-like domain
B. cereus YP_001966639	2.09×10^{-15}	pPER272: 272 kb plasmid, periodontal isolate.	Downstream of a barnase-like domain
Uncultured bacteria YP_133861	3.66×10^{-3}	pRSB101 isolated from sewage.	No neighboring barnase apparent
Deinococcus deserti VCD115 YP_002787535	2.20×10^{-3}	Plasmid 3.	Upstream of a transposase

Table 6.2: Presence of extracellular RNases and inhibitors on plasmids continued

Species	E- value[1]	Location	Notes
Rhodococcus opacus B4 strain YP_002784505	2.69 x 10^{-2}	pKNR-00610	No neighboring barnase apparent.
Acidobacterium sp. MP5ACTX9 YP_004210544	4.15 x 10^{-1}	pACIX903	Annotated as barstar. No neighboring barnase apparent
Butyrivibrio proteoclasticus B316 YP_003832919	1.93 x 10^{-1}	pCY360.	Upstream of a transposase. Found by similarity to bpi.

1.E-values are reported to barnase and barstar unless otherwise specified.

E-value is the chance that a given match to the database could occur by chance, depending on the length of the query sequence and the size of the database (Wit *et al.* 2012). Because E-values are contingent on the specific search query and database, there is no standard threshold for what constitutes a 'good' value. Smaller sequences such as those searched for here inherently give higher E-values, as it is more likely they could occur randomly. Nonetheless, an E-value of 1x 10^{-4} or less is often used for diagnosing a good match (Wit *et al.* 2012).

Only three matches for RNases were larger than this threshold value, and only two for the putative inhibitors. The matches with the higher E-value came from the related plasmids pPER272 and pPER270, with both an RNase and an inhibitor at the same locus, accounting for all but one low E-value. The plasmids were large (~272 and 270 kb, respectively), and have been found in isolates of *B. cereus*. Barnase was annotated as the C-terminal domain on a larger protein with an S-Layer homology (SLH) domain. SLH domains interact with cell wall polymers, tethering enzymes to the surface (Brechtel and Bahl 1999; Sára and Sleytr 2000). Such SLH-domain-containing enzymes have been described in *Thermoanaerobacterium thermosulfurigenes*, including pullulanase (starch degradation), polygalacturonate hydrolase and xylanase (Matuschek *et al.* 1996).

The putative barstar-like protein was closely linked with only a 20 base pair intervening sequence. This sequence, arranged as BN-GGAATAGAGGGATTAAATTT-BS, was used as a query sequence for nucleotide blast (blastn). This search produced more matches on *B. cereus* plasmids, and, in different strains, on *B. cereus* chromo-

somes. Linked chromosomal occurrences of barstar and barnase were also seen in *Rhodococcus* spp., overlapping by one nucleotide. The remaining RNase match with a low E-value, from *B. phymatum*, did not appear to have a linked inhibitor with similarity to any of the search queries.

6.1.2 Testing barstar and secreted barnase as PSK systems in *E. coli*

The ability of secreted barnase to effect a PSK-like activity was tested in *E. coli* strains transformed with plasmids bearing barstar and barnase. Barnase secretion was mediated by the *phoA* signal sequence (Li et al. 1988).

Confirming secretion of barnase

To confirm secretion, the supernatant of induced and uninduced cultures of strain CSBB01 were treated with TCA to precipitate proteins. The resulting protein pellets were resuspended in Laemmli solution and separated using SDS PAGE (Figure 6.1).

Two bands are apparent. The upper band (arrow) at 120 minutes is the size of barnase (~12 kDa) and increases over the time course in induced bacteria but not uninduced bacteria (arrow) (Figure 6.1). The lower band may represent debris from the medium. This finding is evidence that the heterologous *phoA* secretion signal fused to barnase is the cause of an extracellular source of barnase. The medium is also likely to include barnase released from dead cells.

Testing PSK as stability in monoculture

The first set of diagnostics used to tests for PSK activity involved determining of secreted barnase and barstar stabilized the inheritance of a non-conjugative plasmid in monoculture. Plasmids with PSKs are more stable in monoculture than plasmids without (Naito *et al.* 1995; Inglis *et al.* 2013). This PSK test was performed using mid-copy (~20) pBR322-derived plasmids in *E. coli* strain CSH100. Plasmids tested included pBB03, with the genes for barstar and barnase, pBS03, with the gene for only barstar, pLKMCS, a negative control plasmid with only an MCS on the vector backbone, and finally the positive control pLK_par with the type II TA system ParDE. Each plasmid was passaged through serial culturing of the host for 100 generations without antibiotics.

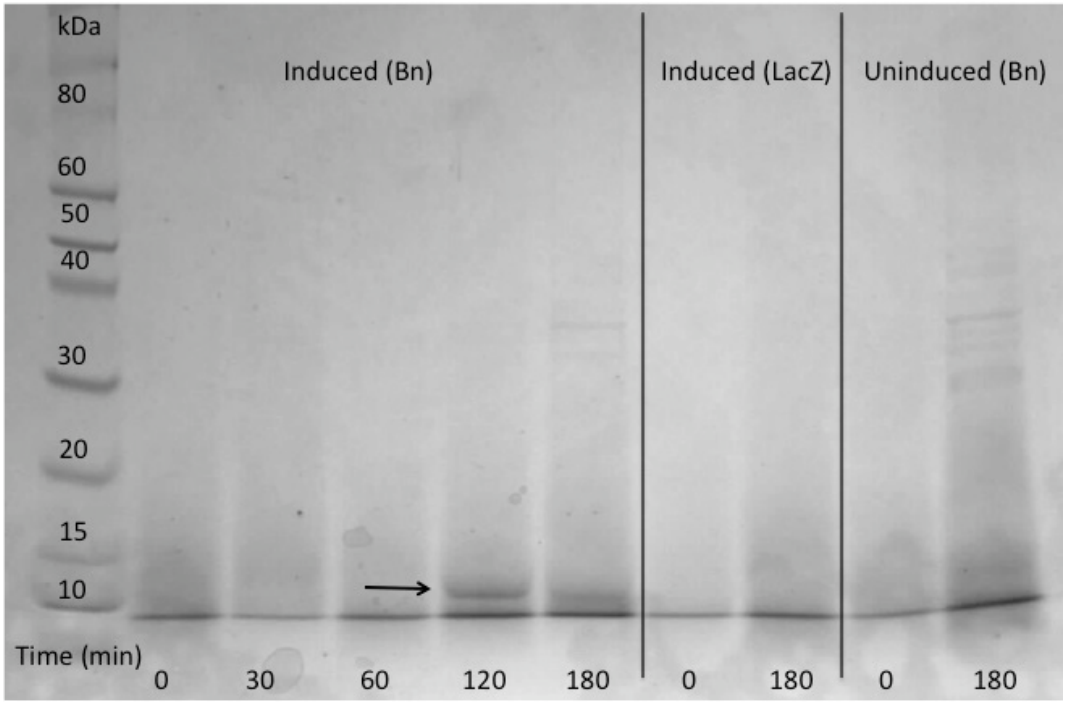


Figure 6.1: Presence of extracellular barnase in the supernatant of CSBB01 *E. coli* strains CSBB01, containing barnase, and CSP01, containing *lacZ*, were cultured to ~0.500 OD_{600} before induction using 1.0 mM IPTG and sampling. Proteins were precipitated from the culture supernatant. Barnase appears in the induced samples at 120 min (arrow)

The pBR322-derived system was stable over 100 generations, with a low dynamic range between the negative (pLK_MCS) and positive (pLK_par) controls. This inherent stability made it difficult to discern between plasmids with or without PSK activity. Interestingly, the barstar-only plasmid was the most unstable, decreasing proportionately to cell density over 100 generations, to ~10% of the total population (Figure 6.3b). Barstar was not induced suggesting that the P_{tac} promoter was sufficiently 'leaky' as to drive barstar transcription, and that or something associated with that expression penalized bacteria with those plasmids. Relative to that plasmid, the barnase (also under the P_{tac} promoter) and barstar plasmid was stable footnoteThese plasmids are not entirely isogenic: the barstar on pBs01 is under the inducible P_{tac} promoter and the barstar on pBB03 is under a constitutive promoter. As it is unable to support full growth of cells containing barnase, barstar is probably expressed at lower levels under P_{tac}., though it was not as stable as the positive control (Figure 6.3b). Taken together, this system did not provide strong evidence for PSK.

The barstar gene alone, or its expression, reduced the growth rate of *E. coli* (Chapter 5), which could cause plasmid-free cells to grow faster in a non-selective medium. I thus hypothesise that barstar has a toxic effect in this heterologous system. While the presence of barnase might be stabilizing the plasmid via a PSK-like mechanism, it is also likely to reduce barstar toxicity by binding *in vivo*.

The second diagnostic for PSK is to test resistance of a plasmid to displacement by another plasmid. These vectors were suitable to transformation-based assays, where a resident plasmid is tested for PSK by transforming the carrier cell with an incompatible plasmid (Naito *et al.* 1995; Nakayama and Kobayashi 1998). Though a common test for PSK, it would not be efficient for systems with an external toxin, as the bacteria are suspended in $CaCl_2$ while being challenged by the incoming plasmid. As expected, there was no evidence that plasmids expressing both barstar and barnase were more competitive under challenge mediated by transformation (results not shown) with an incompatible plasmid. Nonetheless, the resulting doubly transformed cells were useful for additional long-term culturing experiments (Figure C.18), testing if the presence of the incompatible plasmid sped up the loss of the test plasmid (pBB03, pBS01) during serial culturing. This method has been used before to test PSK (Isaeva *et al.* 2010) but was not successful here: instead, the plasmid composition of the cultures appeared to fluctuate in a manner dependent on the antibiotic resistance

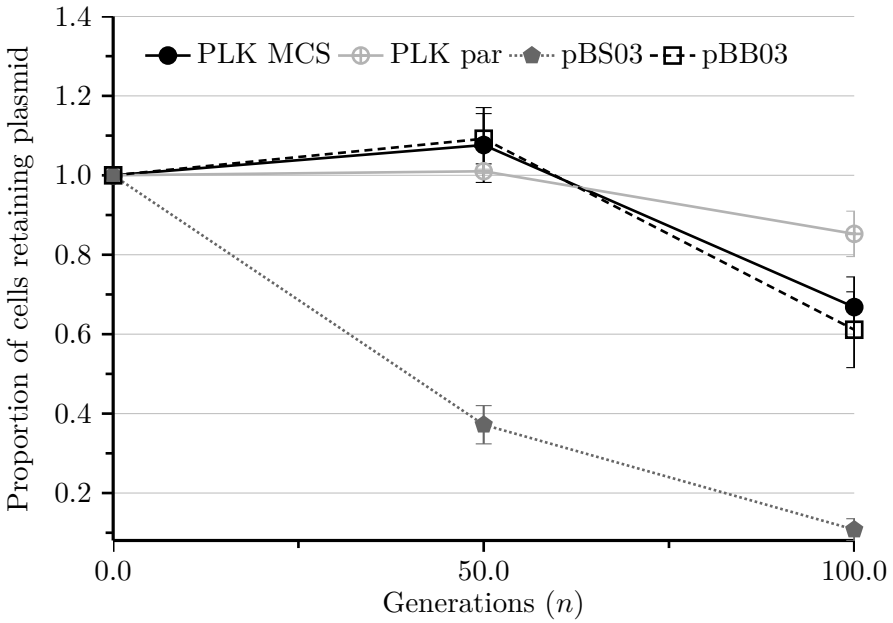


Figure 6.2: Proportion of bacteria retaining barstar and barnase containing plasmids during serial culture without selection. Cultures of *E. coli* strain CSH100 were serially diluted in monoculture without selection for 100 generations, and sampled to determine the number of plasmid-containing cells present. (—●—) Negative control plasmid pLK_MCS. (—⊕—)+PSK plasmid pLK_par. (-⊡-) Test plasmid pBB03. (···⬟···) Test plasmid pBS03. Results are of one trial with four independent replicas of each strain within trial, ± SE.

marker of the plasmid selected for. Previous work in this laboratory has suggested that multiple-copy incompatible plasmids can co-exist in the cell at fluctuating levels depending on the selection pressures (Chapter 5, and Ryan Catchpole, personal communication). This, along with the overall stability of the pBR322 vectors in monoculture, suggests that a lower copy number vector would be a more effective tool for testing PSK.

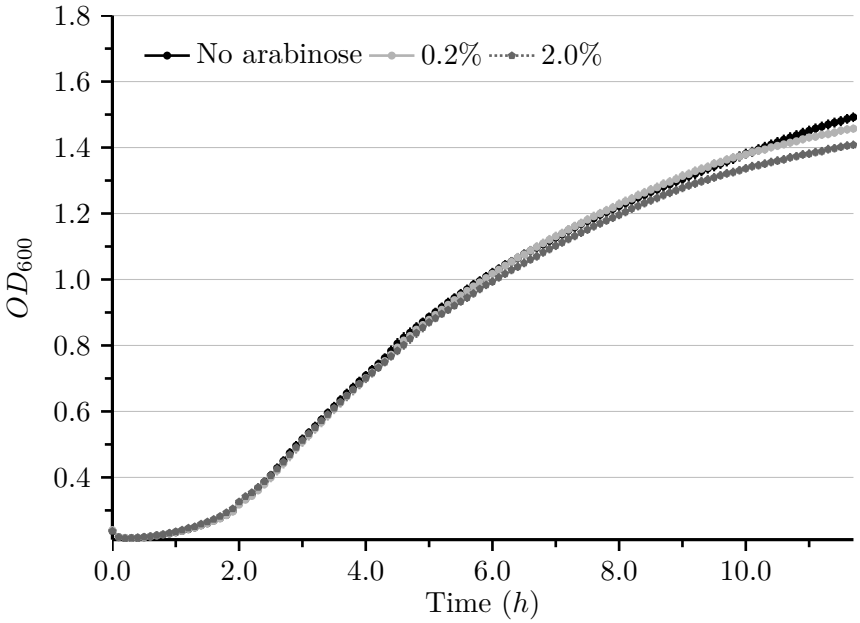
Testing PSK by forcing plasmid loss

A second vector system based on the pHSG415 replicon was devised to test the ability of barstar and barnase to exhibit PSK. The expression system from pBB05 (Chapter 5) was used, where barstar was expressed both constitutively and under the P_{tac} promoter, and barnase was expressed under the PBAD promoter. The co-expression of barstar reduced the toxicity of barnase production (Chapter 5). This expression system was inserted into the low copy number, temperature sensitive vector pHSG415 to give pHS_BB. The plasmid pHSG415 was suitable for three standard PSK tests, including stability in monoculture, transformation with a challenging plasmid, and temperature-induced plasmid loss. Temperature-induced plasmid loss is similar to stability in monoculture, except that the plasmid is lost from the majority of the population within a few hours. This speeds up the assay and reduces the effect that growth of plasmid-free cells is likely to have on the experiment.

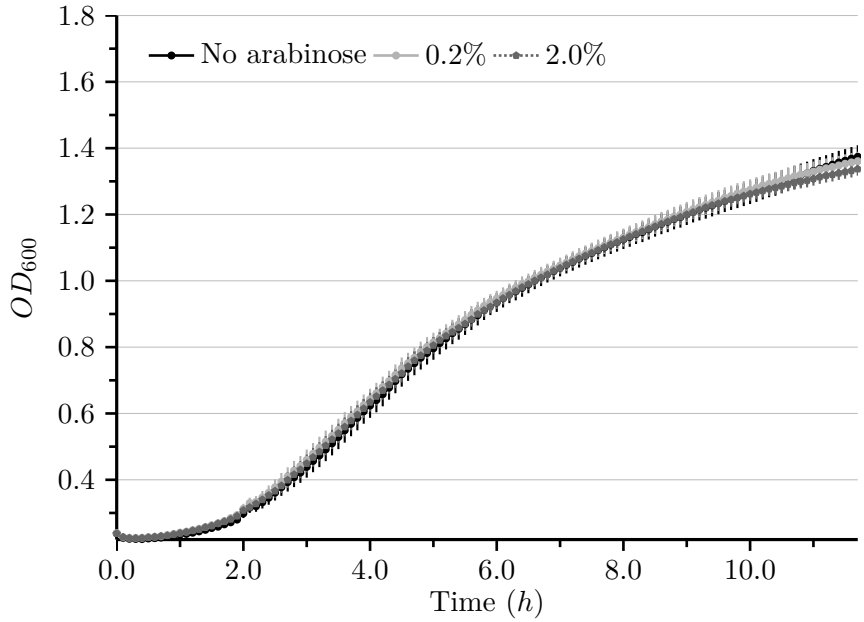
To determine the level of barnase expression that the bacteria could tolerate for subsequent PSK tests, barnase was induced from pHS_BB with 0.2% and 2.0% arabinose (Figure 6.3(C)). Addition of arabinose reduced the population growth of the bacteria harboring pHS_BB, with a significant difference between induced and uninduced at twelve hours (for both, $p \leqslant 0.001$). No significant difference was seen in the positive control (Figure 6.3(A)) and pTN9 (Figure 6.3(B)) in the same time period (for all condtions, $p \geqslant 0.05$).

As the addition of arabinose did not affect growth of the controls, and resulted in only a small decrease in barnase-containing bacteria, future tests were carried out at 2.0% arabinose. Barstar was not induced because constitutive expression was sufficient to protect the cell under this system.

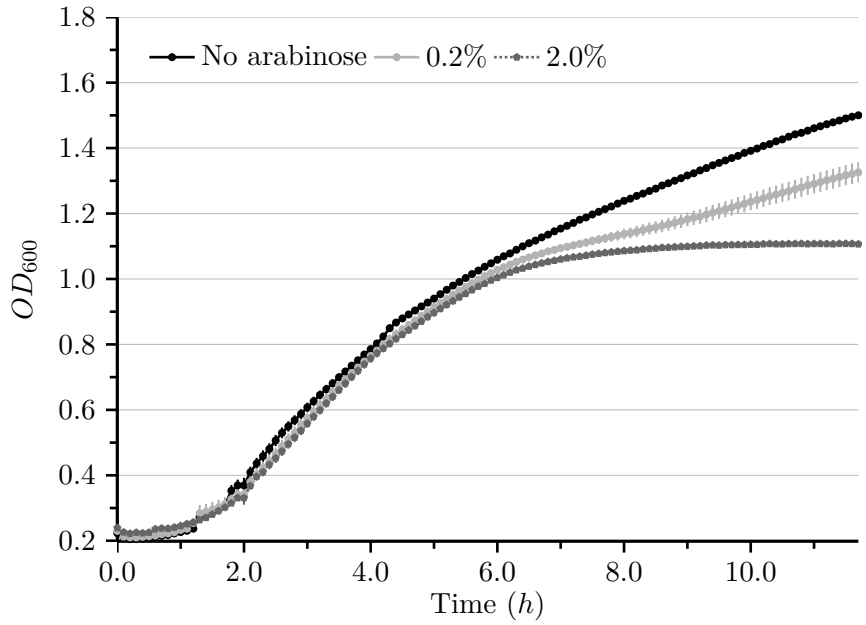
The temperature-sensitive plasmids pHS_BB, pHSG415 and pTN9 were used to test PSK as before (Chapter 4), under inducing (2.0% arabinose) conditions.



(a) pHSG415



(b) pTN9



(c) pHS_BB

Figure 6.3: Growth of *E. coli* bearing temperature-sensitive plasmids, with and without barnase, under barnase-inducing conditions. Cultures of temperature sensitive plasmids in strain CSH104 were grown at 30°Cto 0.1-0.2 OD_{600} before induction with no (—●—), 0.2% (—⊕—) or 2.0% arabinose (···⬟···). (a) Negative control pHSG415. (b) Positive control pTN9 with with known PSK PaeR7. (c) Test plasmid pHS_BB, with barstar and secreted barnase under P_{BAD}. Results are from three trials with four within trial replicas, ± SE.

Cultures were grown at 42℃, diluted as necessary to keep at exponential phase. Plasmid pHSG415 is temperature sensitive due to a mutation in the replication machinery. At 42℃ this function is compromised, eventually leading to dilution of the plasmid by cell division and giving rise to plasmidless daughters. If the plasmid contains a toxin/antitoxin pair that confers PSK activity, those daughter bacteria die.

Barstar and barnase did not confer PSK activity on the plasmid (Figure 6.4). Cultures containing pHS_BB were not significantly different than the no-PSK control at four hours ($p \geqslant 0.05$), where the number of total CFUs in the culture (as measured on LB plates) grows exponentially while the number of CFUs retaining the plasmid (as measured on a selective medium) stays the same. Growth is reported as the log ratio of total cells to plasmid-containing cells over time.

6.1.3 Antibiosis of neighboring bacteria from barnase secretion in *E. coli*

For barstar and barnase to confer a PSK phenotype using an external toxin, barnase would need to be not only efficiently secreted, but also effect death in neighboring bacteria. One group (Ramos *et al.* 2006) demonstrated an effect of barnase on neighboring bacteria using an antibiosis, or zone of inhibition (ZOI) assay. Assays were performed similar to those reported by that group using custom constructs.

Barnase-expressing bacteria were spotted and cultured on plates for 24 hours to allow for secretion of barnase into the medium. Test cultures of mid-log phase *E. coli* were then laid over the spots and incubated for 24-48 hours. Inhibition of lawn growth resembles a 'halo' around the spot and these would be expected if secreted barnase were acting as a toxin.

The first set of experiments was performed using LMBB02, an *E. coli* strain bearing barstar and barnase under control of the P_{BAD} promoter. LMLacZ was used in control spots, and LMG194 (parental strain) was the overlain culture, used at mid-log and stationary phase. Because Ramos *et al.* (2006) showed that barnase had the greatest effect when growth of the bacterial lawn was limited, the assay was repeated with LB, RM, and M9 media at 30 and 37℃. No halos were seen under any of these conditions (Figure 6.5). The effect of barnase expression was apparent in the LMBB02 spots, which grew sparsely on +arabinose plates.

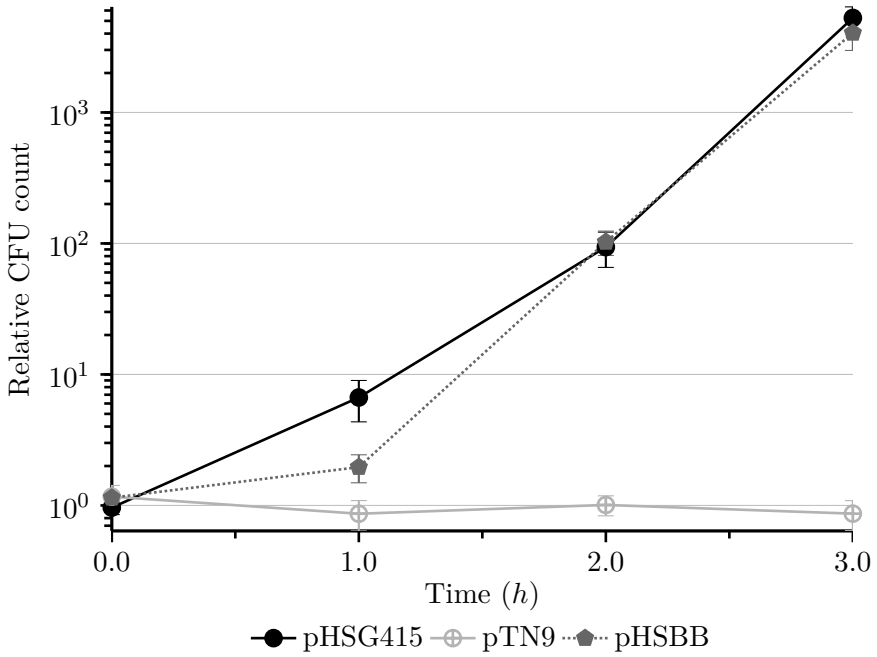


Figure 6.4: Effect of barstar and barnase on bacterial growth under conditions of forced plasmid loss in *E. coli*. *E. coli* strain CSH104 containing potential PSK genes barstar and barnase, on a temperature sensitive (TS) plasmid, were cultured at 42℃to force plasmid loss. The medium was supplemented with 2.0% arabinose. To obtain relative CFU count, number of total live cells (plated on LB) was divided by cells containing the plasmid (plated on antibiotics). Operons inducing PSK would be expected to kill cells losing the plasmid, giving a relative CFU count of ~1. (—●—) TS plasmid pHSG415, no PSK (—⊕—)TS plasmid pTN9 with known PSK, PaeR7. (···⬟···) TS plasmid pHS_BB with barnase, whose expression is induced with arabinose, and barstar. Experiments were repeated in three independent trials, with two replicas per condition per trial, ± SE.

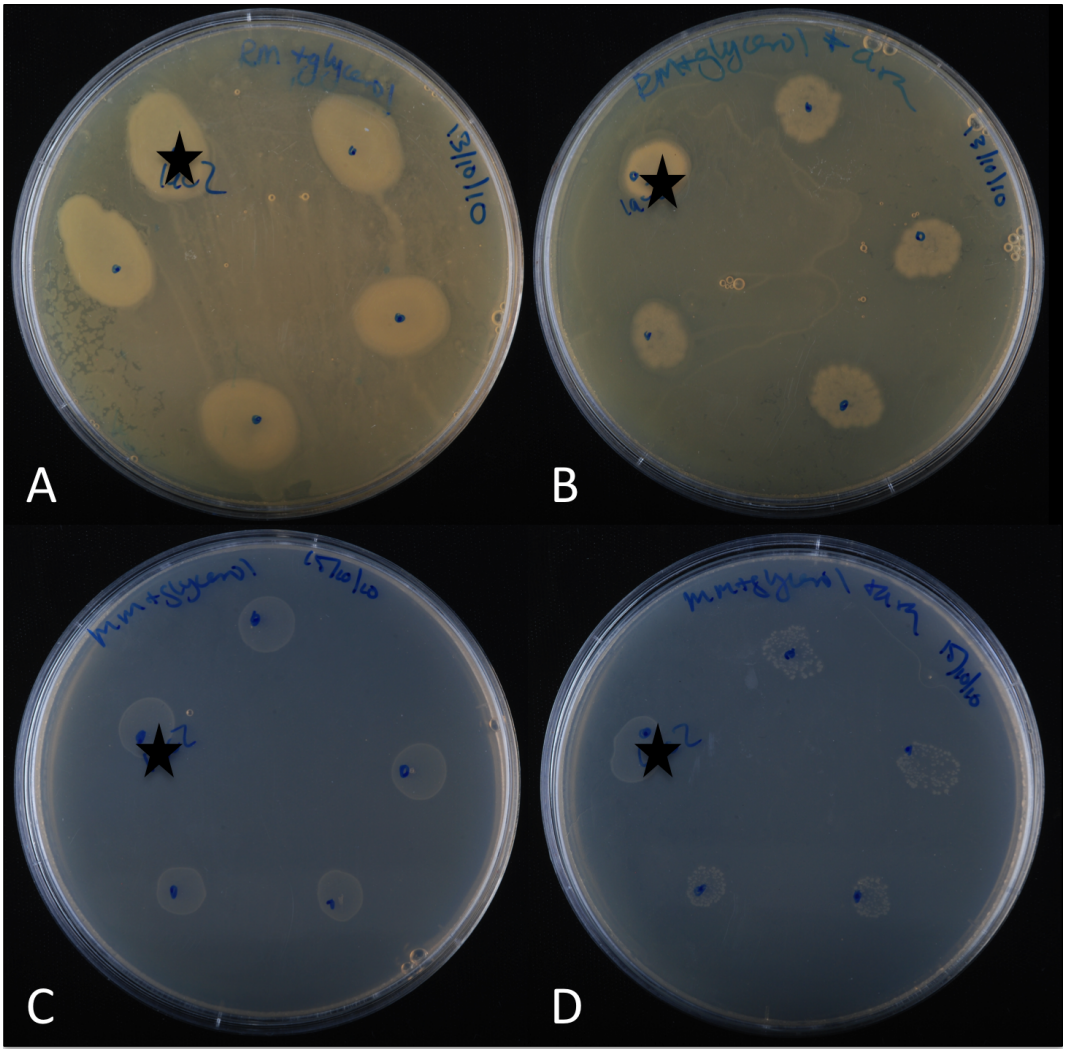


Figure 6.5: Zones of inhibition from *E. coli* secreting barnase from pBB02
E. coli strain LMG194 (lawn) was grown over spots of LMLacZ (star) and barnase-expressing LMBB02. A) RM medium B) RM medium with 0.2% arabinose to induce barnase C) M9 medium B) M9 medium with 0.2% arabinose to induce barnase. This is representative of three trials.

Additional tests were performed by spotting lysed and unlysed cultures of LMBB02, after two hours of barnase expression, directly on lawns of LMG194. Lawns were also spotted with the supernatant of barnase-expressing bacteria, filtered and treated with TCA to precipitate proteins. A total of 64 different conditions were tested using these plasmids, and no halos were apparent.

A final set of experiments was performed using the plasmid pBB05 in strain CSH104 (C4BB05), where both barnase and barstar were induced at the same time, potentially increasing the sub-lethal intracellular concentration of barnase and thus the amount available for export. Cultures of C4BB05 were grown overnight in medium with IPTG to pre-induce barstar expression before spotting on plates with IPTG and arabinose to induce both barnase and barstar. Strain C4Bs01, which also expresses barstar in response to IPTG, was used as a control. These were grown in an incubator for 24 hours before overlaying with CSH104. Again, the trials were performed at various temperatures and using different kinds of media, and no halos were apparent (Figure 6.6). The spots expressing just barstar grew sparsely, similarly to results elsewhere.

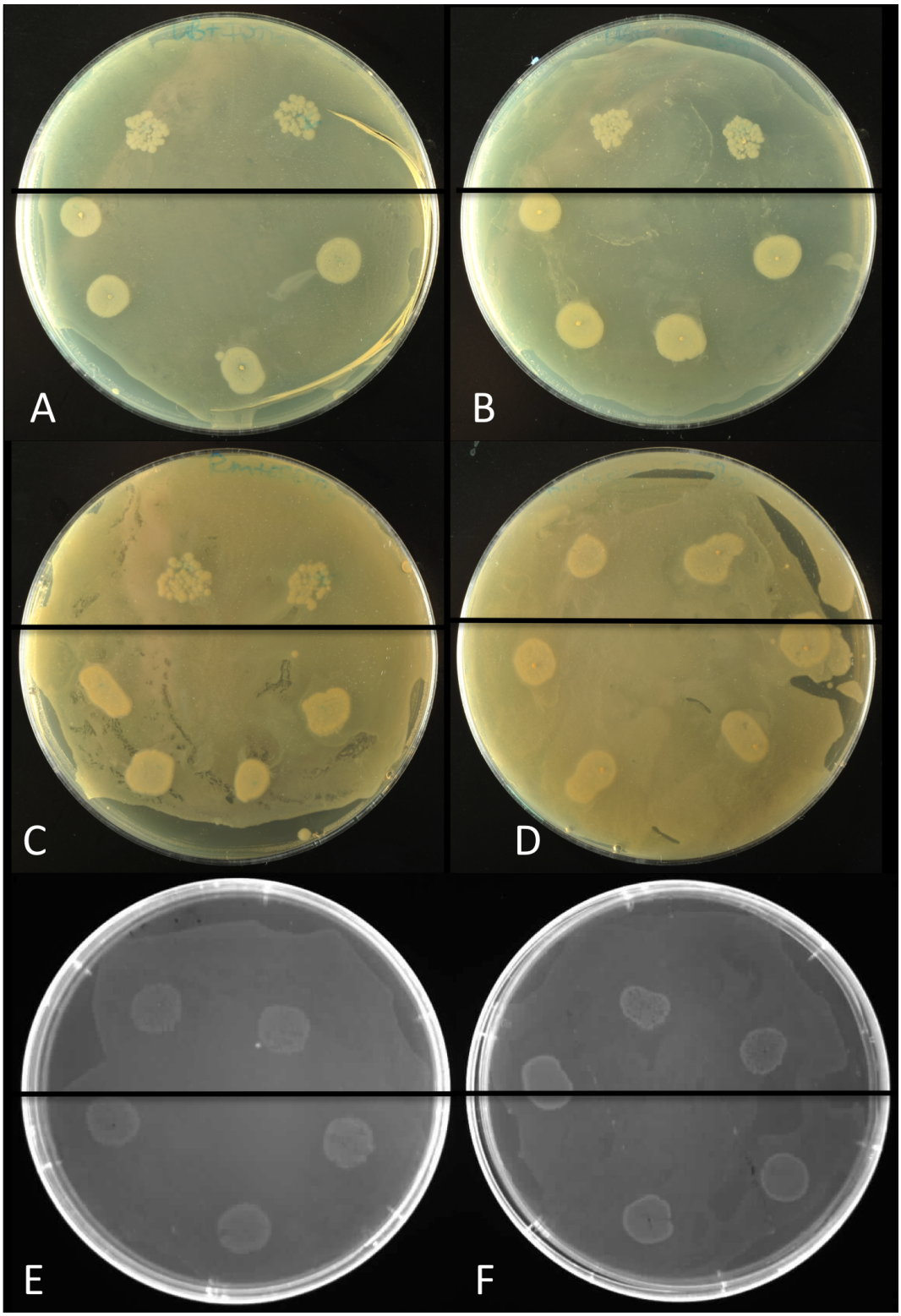


Figure 6.6: Zones of inhibition from *E. coli* secreting barnase from pBB05
E. coli strain CSH104 (lawn) was grown over spots of control strain C4BS01 (above line) and barnase-expressing C4BB05 (below line). A) LB medium B) LB medium C) RM medium D) RM medium with 0.2% arabinose to induce barnase E) M9 medium F) M9 medium with 0.2% arabinose to induce barnase. This is representative of three trials.

6.2 Discussion

Here I investigate the ability of barstar and a secreted barnase to confer a PSK phenotype. They have the standard toxin and antitoxin form of PSK systems, and others had suggested that the genes conferred an enhanced stability on plasmids seen in as few as 36 generations (Ul'yanova *et al.* 2007). PSK and PSK-like mechanisms are known to arise from other external toxins, as in the case of some bacteriocins. The linkage of the immunity and toxin function in bacteriocins allow them to promote their own spread and maintenance in bacteria. I ask if barstar and barnase may act in a similar manner, opening the possibility that by simply changing context (chromosome to plasmid) the genes will confer a PSK phenotype.

6.2.1 Barstar and barnase are not widespread on plasmids and viruses

The distribution of barstar, barnase and related RNases on chromosomes has previously been investigated (Condon and Putzer 2002). Barstar has an entry (PF01337) in Pfam, a database of protein families that includes distribution information based on HMMs. A previous search of genomic sequence data for barnase and barstar on found instances of the genes on chromosomes of divergent species, particularly barstar (Condon and Putzer 2002). I extend this analysis and ask if barstar and barnase had been selected for on plasmids in their native environment.

On the *B. amyloliquefaciens* chromosome, barstar and barnase are not adjacent and are expressed using different promoter systems (Ulyanova *et al.* 2011). According to the operon hypothesis, genes grouped by function have an increased chance of successful transfer and functionality in a new host (Lawrence and Roth 1996). As barstar is necessary to protect the cell from barnase, barnase is less likely to be mobilized on its own. Potential as PSK aside, the separation of the genes on the chromosome does not lend itself to mobilization. But sequence data from one chromosome does not preclude the genes from being organized in an operon and/or mobilized in other strains or species.

Amino acid sequences of barstar and barnase, as well as RNases and inhibitors known to interact with barstar and barnase, were used to find similar matches on plasmids and viruses using blastp in the NCBI database. Our search produced no matches on viral genomes. The strongest match on plasmids was a large, potentially

cell wall-anchored protein with a barnase-like domain, followed by a barstar-like protein. They were seen on the mega plasmids pPER272 and pPER270 from isolates of *B. cereus*. Isolated instances of the proteins in operons on chromosomes were also seen.

While barstar and barnase homologues can be organized as operons and can be on plasmids, no major pattern of mobilization was seen from the data here. Matches were few, and many had a high E-value ($>10^{-4}$). The search done here was preliminary, but more sophisticated computational analyses were deemed unwarranted.

Given the presence of interacting toxins and antitoxins on chromosomes able to act as anti-addiction modules, there may be no selective pressure for mobilization of secreted RNases in their native environment. Like many antibiotic resistance mechanisms, they may have to move to a new environment to become addictive. Using *E. coli* as a model system, with no known interactions with barstar and barnase, I tested their ability to induce PSK in new genetic and cellular contexts.

6.2.2 Secreted barnase is insufficient to drive the PSK phenotype

PSK was first examined via the phenotype of plasmid stability in monoculture. I confirmed that that barnase was successfully transported by the *phoA* signal sequence. It was inserted along with barstar into a pBR322-derived plasmid and cultured without selection for 100 generations.

The control plasmid just containing barstar had a high rate of loss from the system, probably due to the low rate of division of barstar-containing bacteria. This highlights a general difficulty of using stability in monoculture as a diagnostic of PSK activity- if the toxin/antitoxin genes are costly and the toxin not strong (not all bacteria die after loss), plasmidless daughters can escape and increase in number. Though the difficulty here was with barstar only, it would be similarly true if the method were used with an induced toxin that caused residual toxicity over a large number of generations.

Overall, the pBR322 vector system did not exhibit sufficient dynamic range between the positive and negative control to be a good test of PSK. The no-PSK control was not lost rapidly enough from cultures composed of bacteria transformed with either one or two plasmids. The second vector system I used to test PSK of barstar and barnase was pHSG415, a low copy number temperature sensitive system which

allows for forced plasmid loss. Here, barnase was induced with arabinose to increase toxin in the environment. While this reduced growth over 12 hours it had no effect on growth relative to the negative control within the four hour time period of the PSK assay.

Using this system, I found no indication that barstar and barnase were conferring PSK. It is possible that the rapid growth conditions and frequent dilutions used during the temperature-sensitive assay diluted the secreted barnase to a point where it could not longer harm neighbouring bacteria. To drive this phenotype, the secreted toxin would need to be expressed at high enough levels to cause harm to neighbouring bacteria. Only one published experiment has shown this of barnase (Ramos *et al.* 2006). I sought to replicate this experiment using my constructs, to confirm that barnase is toxic to non-producing cells and elucidate the conditions under which it is so.

6.2.3 Secreted barnase does not cause ZOIs within this test system

Ramos *et al.* (2006) reported that *E. coli*, soil bacteria and yeast were susceptible to barnase secreted from *E. coli* with the *phoA* signal sequence, under some growth conditions. Using similar methods of growing barnase-secreting strains on plates before overlaying with a non-producing strain, I screened for halos, which are indicative of zones of inhibition, in a wide range of conditions, varying expression system, medium, inducer concentration, and temperature. Lawns were also spotted with lysates and concentrated supernatants from producing bacteria. Under none of these conditions were ZOIs observed.

Ramos *et al.* (2006) used a different expression system (Ramos *et al.* 2005) than the one tested here. The plasmids pHRBar-6 and pHRBar-13 encode barnase under the P_{tac} promoter, and two barstars: one constitutive and one under the *nifA* promoter, inducible with the NifA protein. Induction of barnase with IPTG from these constructs causes cell death, with some growth seen in strains producing NifA to co-induce barstar (Ramos *et al.* 2005). ZOIs were reported when barnase was uninduced, with larger zones when barnase was induced in the presence of NifA (Ramos *et al.* 2006).

This is most similar to the construct pBB05, which encodes a constitutive and inducible barstar and an inducible barnase. Barnase in my system is under the P_{BAD}

promoter, which does not result in as much barnase production as the P_{tac} promoter when induced. This may mean that I was not able to produce enough barnase in my system, though it seems unlikely that uninduced P_{tac} expression would be greater than induced P_{BAD} expression. I was unable to obtain their constructs for this work.

The next obvious step in this process would be to confirm that barnase was bacteriocidal when applied externally to bacteria using purified protein. This would also provide a starting point for deciding the quantity of secreted barnase necessary for PSK. Bacteriocins and their inhibitors are able to exhibit PSK with only small amounts of toxin in the medium (Inglis *et al.* 2013), and it would be interesting to investigate the minimal amount of growth inhibition of neighbouring bacteria required for PSK to be effective.

Barnase has been proposed as an agent of bacterial competition (Ulyanova *et al.* 2011), and it might still be toxic as secreted from *B. amyloliquefaciens*. Indeed, a common explanation for the presence of solitary barstar genes on chromosomes is protection against barnase secretion from other bacteria in the environment (Condon and Putzer 2002). But the lack of apparent toxicity of secreted barnase under the conditions tested here would explanation the lack of PSK activity.

Despite this, it is still valid to consider the general possibility that secreted toxins are able to drive replicon competition, and thus be selected for on plasmids. An externally-driven PSK would have different population dynamics than one relying on an cytoplasmic toxin. For both types, any cell losing the plasmid from a homogenous population of plasmid-bearers would be killed, including competitors that facilitated the loss. But if the plasmid was invading a population, the external toxin could also eliminate potential hosts. Bacteriocins are often induced under conditions of stress (Cascales *et al.* 2007), when population density is high. This may reduce the chance that potential hosts are harmed before the plasmid has a chance to invade. It also creates a system of conditional PSK, driving retention during times of stress.

It could that PSK can be generated simply by a change in replicon: the movement of the genes for a secreted toxin and its antitoxin to a plasmid. It would be expected that such genes would be selected for maintenance on plasmids once acquired. Known type II TA systems often occur in genomic regions near potential secretion systems (Benz and Meinhart 2014; Guglielmini and Van Melderen 2012), though none have been shown to act through translocated toxins yet. Recent computational analysis

of *E. coli* gut metagenome data has suggested that secreted proteins are more likely to be mobilized (Nogueira *et al.* 2009), particularly virulence factors. While this has been used to support the infectious relatedness model (Chapter 1), it is possible that some of those were linked secreted toxins and antitoxins able to drive their own maintenance.

Chapter 7

Conclusion

A rich diversity of genetic elements make up the prokaryotic gene pool, capable of reproducing vertically and horizontally between bacteria. Changes in a gene's environment as it moves between cells and within and between replicons can alter the gene's phenotype. I am interested in what affects the reproductive success of genes as they move into new genetic and cellular contexts.

The distribution of genes on what are called chromosomes and mobile genetic elements is not random, with some genes appearing to differentially accumulate on one type of replicon. Genes that exhibit PSK are particularly successful on mobile elements. All PSKs identified thus far have been toxin and antitoxin gene pairs, including the canonical TA systems, RM systems, bacteriocins, and abortive infection systems. But there is considerable diversity within these types: some types of toxin and antitoxin gene pairs do not induce PSK, and within types that commonly do, some families and homologues do not. In this thesis, I investigated the overlap between the structure and biochemical functions of TA systems and the phenotype of addiction

7.1 PSK is difficult to predict from gene structure and protein biochemistry

Our ability to identify putative new genes and gene families has increased in conjunction with the upsurge of genome sequencing, augmented by the development of new computational methods by which to analyze these data. The computational methods used to find new TA systems frequently rely on the sequences and structure of experimentally verified TAs. We used similar methods to look at the distribution

of type I TAs, which are apparently distributed across less taxa and on fewer horizontal replicons than type II systems (Chapter 3). Our ability to infer function from these methods was tested more specifically with the system PT-ptaRNA1, which structurally resembled type I TA systems, with a predicted membrane protein and an antisense RNA transcribed from the opposite strand (Findeiß *et al.* 2010). Loci of the family have a distribution consistent with spread by HGT (Findeiß *et al.* 2010). Despite this, expression of the predicted toxin ORFs did not reduce growth (as measured by saturation density) in *E. coli*, and the operon did not induce PSK (Chapter 4). Findings such as this highlight that structural and distributional similarity of genes is not sufficient to predict PSK systems without experimental validation. It would also support the value of non-structure based searches, such as the identification of regions intractable to cloning that are likely to be toxin-bearing, as was done by Sberro *et al.* (2013).

PT-ptaRNA1, though structurally and distributionally similar to type I TAs that induce PSK, did not have the biochemistry necessary to exhibit PSK in *E. coli*. Using biochemisty as a starting a point, I looked at a gene pair, barnase and barstar, known to exhibit a toxin and antitoxin phenotype and asked what would be necessary for them to exhibit PSK, in a manner akin to both type II TA systems, with a cytoplasmic toxin (Chapter 5), and bacteriocins, which have a secreted toxin (Chapter 6). The toxin and antitoxin functionality was also insufficient to predict PSK. Specific amounts of protein expression and stability are necessary for an cytoplasmic toxin to induce PSK. Secreted toxins must be efficiently taken up by neighboring cells. Specifically for our system, barstar lacked the unstable regions often seen in labile antitoxins (Chapter 5) (Lubienski *et al.* 1994)- stable antitoxins require close titrating to get PSK after plasmid loss, which can be difficult to engineer (Chapter 5). Barnase was both very toxic when expressed internally, and yet not toxic to neighboring cells when expressed from these constructs (Chapter 5, 6). I drew the conclusion that these biochemical characteristics are also insufficient evidence for reliably predicting PSK.

7.2 PSKs are context dependent

Structural similarities and biochemical similarities are not sufficient to determine whether a given system will act as a PSK because numerous contextual factors have

an effect on whether the genes are addictive. A given set of genes may have the phenotype in one species but not another, under one set of environmental conditions but not another, or on one replicon but not another.

This differential PSK expression is evident from analyses of known TA systems. Kawano *et al.* (2002) hypothesized that the lack of PSK induction by Lrd-RdlD in their host was due to under-expression. And while recent studies put Ldr and Fst in the same superfamily (Fozo *et al.* 2010), they have disparate phenotypes in their respective hosts. Ldr occurs as chromosomal repeats in Gram-negative bacteria, and does not exhibit PSK when placed on a plasmid (Kawano *et al.* 2002; Kawano *et al.* 2007). Fst is found on plasmids and exhibits PSK in Gram-positive bacteria (Weaver *et al.* 2009).

Even within a given bacterium, the cellular environment is neither uniform nor static: perturbations in the expression of PSK genes and their rate of decay can change their effectiveness. This is evident with some ABIs and chromosomal TAs, which can be activated during changes in global transcription and translation rates (Engelberg-Kulka *et al.* 2006; Amitai *et al.* 2004; Christensen *et al.* 2003; Cheng *et al.* 2004; Engelberg–Kulka *et al.* 1998; Slavcev and Hayes 2003), and increase in protease levels associated with stress response (Christensen *et al.* 2004; Gerdes 2000). Work described in this thesis would suggest that a decrease in the rate of cell division would also affect PSK by effectively increasing the toxin half-life (Chapter 5).

Genes on mobile elements are potentially exposed to an even greater variety of cellular environments as they move into different hosts. Many TAs are no longer addictive in strains that do not have the right proteases (Brzozowska and Zielenkiewicz 2013; Dao-Thi *et al.* 2000; Van Melderen *et al.* 1996), for example. Weaver *et al.* (2009) found they were able to clone only one of three *fst-rnaII* loci from *E. faecalis* into *E. coli*, potentially due to over-expression of the toxin in the novel host. This can make the PSK phenotype difficult to validate outside of the hosts that they are found in.

7.3 Some PSKs may be more efficient across different taxa than others

Such context dependence may offer an explanation for the apparent diversity of mobility across TA families. Data on validated TAs suggests a correlation between distribution, mobility and the PSK phenotype (Chapter 3) (Leplae *et al.* 2011; Fozo *et al.* 2010; Makarova *et al.* 2009). A similar pattern emerges between different types of RM systems (Mruk and Kobayashi 2014). All TAs that can induce PSK are likely to be successful on mobile elements, but may exhibit differential success during periods of environmental change, internal or external to the cell, and while moving into new hosts. This could in turn be reflected in differences in phylogeny (Chapter 3, 5).

The toxin target has been speculated to cause some TAs to be more successful across taxa (Leplae *et al.* 2011; Goeders and Van Melderen 2014). PSK requires an effective toxin: as such, the phenotype is most likely to be retained across divergent species if the toxin targets highly conserved aspects of cellular function. The movement of the module into a new host may be unsuccessful if the antitoxin exerted a toxic effect on the cell. This was seen here with barstar (Chapter 5) and was proposed for type I RNA antitoxins (Chapter 3), both of which might bind off-target substrates (Jackson *et al.* 2003).

The distribution of TA systems could be compared to the conservation of the target of a given family's toxin. Many type II TAs target translational machinery, which is highly conserved across all domains of life (Poole and Logan 2005; Noller 2004). DNA synthesis machinery, DNA polymerases and primases are less conserved (Poole and Logan 2005; Leipe *et al.* 1999; Werner and Grohmann 2011; Aravind and Koonin 2001). Amongst type I TAs, SymE, which targets mRNA (Kawano *et al.* 2007), would provide an interesting comparison to the membrane proteins investigated within this thesis.

Clearly, neither toxin nor antitoxin target accounts for all differences within TA systems, because families with the same target often contain loci that differ in ability to confer a PSK effect on a plasmid (Szekeres *et al.* 2007; Christensen *et al.* 2004; De Bast *et al.* 2008; Wilbaux *et al.* 2007; Fiebig *et al.* 2010). Similarly, type I and III RMs are not as mobile and have not been shown to induce PSK (Naderer *et al.* 2002; O'Sullivan *et al.* 2000; Mruk and Kobayashi 2014), but have the same target (DNA)

as type II RMs. Other factors are also important for PSK, including the levels at which the toxin and antitoxin are expressed, and the rate at which the toxin and antitoxin are degraded in the cell after plasmid loss (Chapter 5).

TA systems rely on differential decay of the toxin and antitoxin to induce PSK (Chapter 5). Type II systems are tightly regulated by protein antitoxins, which interact with the toxin and act as transcriptional repressors, sensitive to changes in stoichiometry (Mruk and Kobayashi 2014; Zhang *et al.* 2003; Kedzierska *et al.* 2007; Cataudella *et al.* 2013; Cataudella *et al.* 2012; Afif *et al.* 2001). I used equations of logarithmic decay as a starting point for analyzing the conditions necessary for type II TAs to exhibit PSK. In particular, I analyzed the population of toxin and antitoxin in the cell necessary for PSK given their respective half-lives (Chapter 5). Antitoxins with similar stabilities as their associated toxins cannot be expressed in numbers significantly higher than the toxin, or there will not be sufficient free toxin to create a PSK effect within a given time period. Given the inherent noise of gene expression (Raser and O'Shea 2005; Rao *et al.* 2002), systems relying on stable antitoxins risk plasmid suicide from excess toxin in the cell prior to loss. Thus, there is likely to be an evolutionary trend toward TA systems with high expression levels of very unstable antitoxins.

Both type I and type II TA systems tested so far express large populations of antitoxin relative to the toxin, with subsequent rapid decay (Van Melderen *et al.* 1996; Franch *et al.* 1997; Aizenman *et al.* 1996; Cataudella *et al.* 2013; Sayed *et al.* 2012). The type II TA antitoxins MazE and RelE are 8 and 10-fold less stable than their toxins, respectively (Aizenman *et al.* 1996; Cataudella *et al.* 2013), and the type I TA antitoxins Sok and $\mathrm{SprA1_{AS}}$ are 20 and 40-fold less stable, respectively. RelE is present in the cell in 10-fold excess of the RelB toxin (Cataudella *et al.* 2012; Cataudella *et al.* 2013) which, given the rate of decay, means it is probably expressed 100-fold higher; $\mathrm{SprA1_{AS}}$ is expressed in 30-80 fold excess of the toxin (Sayed *et al.* 2012).

The role that stability and protein population play in PSK has only been examined as general trends, and experimental validation is needed. What ratio of antitoxin to toxin is sufficient to avoid cell death while the plasmid is retained? How stable is a 'stable' antitoxin? This could be tested using synthetic constructs, as was attempted here, though the necessary control over relative expression is difficult to engineer. Additional methods could involve known PSKs, by mutagenesis of regulatory regions

or antitoxin residues known to contribute to stability. The ratio of toxin to antitoxin from known PSKs could be manipulated in trans by titrating out antitoxin with inactivated toxins (Tripathi *et al.* 2012).

This research could also exploit differences in naturally occurring TA systems. The ratio of toxin and antitoxin and their stability in the cell could be measured and compared across TAs that differ in their location on chromosomes and plasmids, and their ability to exhibit PSK. A good data set for this might be the analysis of the MazF toxin, done by Chopra *et al.* (2013). They identified a number of closely related homologues on plasmids and chromosomes, including some subgroups that did not appear on plasmids at all. The relative expression of the toxins and antitoxins in theses systems and their ability to induce PSK would shed light on the proposals here about what makes PSKs successful on mobile elements, and the necessary levels of toxin and antitoxin to achieve PSK.

These factors may explain variation between and within TA systems in exhibition of PSK. Differences between type I and type II systems could be due to common toxin targets and the different types of regulation (RNA versus protein). Differences between individuals of the same family could be due to subtle changes in their relative expression and stability. If the ability to induce PSK affects mobility, these factors could explain differences between the distribution of TA system families on horizontal replicons and the degree to which their inheritance appears to be lineage dependent.

7.4 The competition model supports this account of PSKs

The questions addressed in this thesis were initially framed under the competition model. Competition predicts that genes will be mobilized due to their ability to increase reproduction rates by infectious transfer. This is dependent on whether the genes are advantageous or not to a replicon using horizontal gene transfer in a given environment, including their ability to effect competition between the replicon on which they reside and other replicons.

This mediation of competition explains the prevalence of gene modules such as TA systems, RM systems, and ABI systems on mobile elements. TA systems affect

competition between incompatible plasmids (Cooper and Heinemann 2000; Cooper and Heinemann 2005) but also between plasmids and phage. There is evidence of an 'arms race' between the two (Goeders and Van Melderen 2014): phage T4 encodes a protein that directly inhibits the activity of toxins from multiple TAs able to induce ABI (Otsuka and Yonesaki 2012) and some phage express a protein that inhibits Lon or Clp proteases, stabilizing antitoxins and thus limiting free toxin (Sberro *et al.* 2013; Engelberg–Kulka *et al.* 1998). Phage that escaped ToxIN-induced ABI in the laboratory were found to have expanded repeat regions in their genomes that mimicked the ToxI antitoxin or to have picked up a copy of the entire ToxIN locus via recombination (Blower *et al.* 2012).

Competition between chromosomes and mobile TA systems is seen in anti-addiction modules. A chromosomal *ccd* locus in *Erwinia chrysanthemi* does not induce PSK when on a plasmid, but still produces an antitoxin capable of preventing PSK by the F plasmid *ccd*, conferring a selective advantage over isogenic strains without the chromosomal locus (Saavedra De Bast *et al.* 2008).

The contribution of the competition model for analyzing PSK is more subtle than simply suggesting genes that engage in competition will be selected for on mobile elements. Competition differs from gene-deterministic models of PSK evolution, which expect to predict which genes are more horizontally mobile by the activity of the genes themselves (or at least the known activities). Competition bases prediction on the environment; when an environment creates conditions by which combinations of genes reproduce faster by horizontal rather than vertical transmission (e.g., because they develop a PSK effect), then those genes will differentially accumulate by HGT.

The importance of the environment, as emphasized by the competition model, is illustrated by the evolution of antibiotic resistance. In the presence of lethal concentrations of antibiotics in the environment, plasmids with resistance genes displace related plasmids without them by competition and then spread quickly through populations. Sometimes this involves the mobilization of a chromosomal resistance gene that comes from bacteria that produce the antibiotic, or live in the same environment as the producer strain (Laskaris *et al.* 2010; Benveniste and Davies 1973; Davies and Davies 2010). Resistance may be a result of pleiotrophy and arise from single or multiple genes displaying an effect that was not predicted from their known biochemistry (Davies and Davies 2010; Forsberg *et al.* 2012; Martinez 2012; Laskaris *et al.* 2010).

Thus it is difficult to predict which genes will confer resistance from their original biochemical and genetic context, and it will only be addictive in the right environment. Genes that confer resistance can serve other functions in the cell, from DNA damage repair to efflux pumps, and resistance to one antibiotic can confer resistance to others (Heinemann *et al.* 2000). Native DNA repair enzymes of *Dictyostelium discoideum* confer high levels of resistance to bleomycin and UV light (Deering *et al.* 1996), both of which damage DNA. Mobilization of genes that perform alternative or additional functions in thier host is an important source of resistance mechanisms, relevant to synthetic antibiotics such as quinolines (Poirel *et al.* 2005). Gene deterministic models fail to account for mobilization of these genes, but competition predicts mobilization whether the gene was a resistance mechanism before mobilization or not.

The emphasis of the competition model on gene environment is consistent with the results presented here on PSK. There is no inherent property of a toxin and antitoxin gene pair that allows us to predict whether it will exhibit PSK. It is highly contingent on the environment of the genes. A gene pair that exhibits PSK in a given environment will be selected for on mobile elements.

7.5 Responsiveness of TA systems to stress could be selected for on mobile loci

The ability of TAs to induce PSK may not be the only factor that accounts for their selection on MGEs. The responsiveness of TA systems to cellular stress has been extensively studied for chromosomal TA systems: I speculate here they may also be selected for on mobile TAs. Stress increases competition among replicons by reducing resources and increasing the frequency at which potentially competitive HMEs reside in the same cell. Stress that results in slower cell division may also directly increase the effectiveness of PSK by increasing the half-life of the toxin (Chapter 5). TAs can provide an advantage to mobile elements independent of their ability to induce PSK (Pimentel *et al.* 2005) and enhancing the stress response by inducing a bacteriostatic state could be one such mechanism. TA systems on MGEs can respond to stress stimuli in a manner analogous to chromosomal TAs. For example, small increases in CcdB toxin from the F plasmid resulted in a substantial (150-850 fold) increase in persister cell formation in the presence of five antibiotics (Tripathi *et al.* 2012). The

Ccd TA system, then, is a transferable persistence factor.

Gene transfer is known to increase during times of cellular stress (Heinemann 1999). These stress signals can involve nutrient limitation, stationary growth phase, and quorum sensing. Competence necessary for natural transformation can be signaled by quorum sensing peptides (*B. Subtilis* and *Streptococcus pneumoniae* or nutrient starvation (*Haemophilus influenzae*) (Lorenz and Wackernagel 1994) or DNA damage (Claverys *et al.* 2006; Prudhomme *et al.* 2006; Claverys *et al.* 2009). The F plasmid exhibits a form of entry exclusion that reduces entry of addition F plasmids (Heinemann and Ankenbauer 1993), which is relaxed during stationary phase (Peters and Benson 1995).

The presence of antimicrobial compounds can increase gene transfer. Competence in some Gram-positive bacteria is also induced in response to antibiotics (Claverys *et al.* 2006; Prudhomme *et al.* 2006; Claverys *et al.* 2009). Treatment with the antibiotic ciprofloxacin increases mobilization of ICEs by inducing SOS response in *E. coli* (Beaber *et al.* 2004), and increased recombination of divergent sequences into chromosomes in an SOS-independent manner (López *et al.* 2007). Antibiotics have also been shown to increase transfer of transposons (Heinemann 1999), and increase both the replication and subsequent transfer of pathogenecity islands flanked by integrase genes and repeat elements in *Staphylococci* spp. (Úbeda *et al.* 2005; Maiques *et al.* 2006).

Stresses that reduce vertical transmission, by slowing cell division, do not necessarily decrease horizontal transfer. Conjugation can still occur after treatment with a number of antimicrobial agents, including mytomycin-C and UV radiation (Heinemann and Ankenbauer 1993), depending on the order of the stresses encountered (Heinemann 1999). Furthermore, mathematical modeling of bacterial populations exposed to different levels of stressors (such as some antibiotics) indicated that that under these conditions plasmids proliferate among phenotypically 'dead' bacteria by horizontal gene transfer (Willms et al. 2006).

If the stress would eventually lead to a state that prevents gene transfer, there would be an advantage to slow down growth sufficiently to give the mobile element the chance to move into a new cell. This would be the case for bacteriostatic and some bacteriocidal TAs, but not for modules like Hok-Sok, which drain the membrane potential, and stop conjugation (Heinemann and Ankenbauer 1993).

It is speculated here that induction of bacteriostasis by some plasmid-borne TAs may provide a selective advantage to the replicon, both by mediating competition with incoming elements and potentially giving the plasmid time to transfer into new cells. Studies of chromosomal TA-deletion mutants competed under stressful conditions have not shown that the TAs provide a selective advantage to those cells (Tsilibaris *et al.* 2007). It would be of interest to repeat these experiments under conditions of plasmid competition. Genes exhibiting PSK have been shown to advantage plasmids when competed with other plasmids (Naito *et al.* 1995; Cooper and Heinemann 2000; Cooper and Heinemann 2005) under conditions permissive to conjugative transfer. These experiments could be repeated under media conditions eliciting cellular stress, as in Tsilbaris *et al.* (2007), such as rifampicin treatment, acid shock, and nutrient limitation.

7.6 Conclusion

In this thesis I provide the groundwork for analyzing the necessary conditions for PSKs, an emergent phenotype of some toxin and antitoxin gene pairs. PSK requires specific, highly regulated expression of particular genes in the right cellular and environmental context. But their ability to mediate horizontal competition means that the genes will be highly mobilized in environments where this occurs.

Understanding the necessary conditions for PSK is important in light of current use of these systems in the design of new therapies and technologies. Knowledge of these conditions has been utilized in the design of stable vectors for antibiotic-free protein expression systems (Isaeva *et al.* 2010). Perturbation of TAs has been a suggested target for antimicrobial agents (Williams and Hergenrother 2012), with the hope that activation of the systems would cause stasis or death. Yet recent work on the ParDE/KisKid type II TA system has suggested this strategy has serious flaws for at least some TAs: perturbations in the toxin and antitoxin levels of a ParDE operon caused plasmids, which often carry antimicrobial determinants, in the host to *increase* in copy number (Pimentel *et al.* 2005; Pimentel *et al.* 2014). The authors suggest a greater understanding of individual TA systems would be necessary before targeted for therapeutic purposes (Pimentel *et al.* 2014).

Understanding PSK also increases our knowledge of what makes genes mobile, an important consideration when gene transfer is not desired. The rapid spread of

antibiotic resistance genes is well documented (WHO 2014) and to be expected given the high use of antibiotics (Normark and Normark 2002; Heinemann 1999; Smith *et al.* 2004). Predicting whether genes will transfer out of a host organism is also an important facet in the risk assessment of genetically modified organisms (Heinemann and Traavik 2004a; Ellstrand 2001). Toxin and antitoxin gene systems have been proposed for use in containment strategies for genetically modified organisms. This includes male sterility in plants, which commonly utilizes barnase (García-Sogo *et al.* 2010; Wei *et al.* 2007; Bisht *et al.* 2007; Mariani *et al.* 1990), and conditional lethality in bacteria designed for bioremediation and field release (Davison 2005; Heinemann 2007).

In bacteria, methods have been created to reduce HGT of transgenes by linking them with toxin genes, rendering receipt of the recombinant material lethal (Recorbet *et al.* 1993; Knudsen and Karlstrom 1991; Torres *et al.* 2003). The toxin in many of these constructs come from PSK systems (Knudsen and Karlstrom 1991; Torres *et al.* 2003). Torres *et al.* (2003) proposes a system using colicin E3 and the endonuclease from EcoR1 linked to a transgene on a plasmid, with the immunity functions on the chromosome. Given the ability of genes to transfer from chromosome to plasmid, it is quite plausible that this system designed to reduce HGT from bacteria would, ironically, instead serve to increase it (Heinemann and Traavik 2004b).

Significant work has been done on chromosomal TAs, including details of their regulation and stability, and how this affects stress-response in the cell. But little has been done on TA systems on mobile elements, the conditions necessary for PSK, and the relationship between their distribution across replicons and phyla. Much of this work is by necessity preliminary, but gives us the framework from which to continue addressing these questions. These analyses contribute to our knowledge of the movement and evolution of PSK systems and gene mobility as a whole.

This work also addresses the issue of the 'one gene one phenotype' hypotheses.

Early assumptions of genetics assumed that a given gene, or allele, would consistently produce one polypeptide and one phenotype <Prasun et al. 2007>. This belief is still pervasive in modern genetics, despite the accumulation of evidence to the contrary.

In fact, the same level of contextualization discussed here for PSK is necessary for the expression of many phenotypes. The final protein produced from a gene can be altered during changes in transcription, RNA processing, translation, and

post-translation modification (Prasun *et al.* 2007). Effect of this protein on the cell is then mediated by the presence or absence of inhibitors, substrates, and proteases, which are also subject to the production processes listed above.

Recent advances in human genome sequencing, for example, have not provided the simple gene to disease associations once assumed to exist (Cooper and Shendure 2011). For example, a similar genetic lesion in the human beta globin gene causes sickle cell anemia, but with the disease outcome ranging from life-threatening to symptom-free (Quinn and Miller 2004).

An increase in understanding of the role of context will aid in our attempts to predict complex phenotypes, as well as highlight those areas that defy simple deterministic assumptions.

Appendices

Appendix A

Formulas and Solutions

A.1 Media

Plates made with 16g/L of bacteriological agar (Oxoid).

A.1.1 Luria Burtani Medium

20 g/L LB Broth Base (Invitrogen)

A.1.2 RM medium

1x M9 broth (see below)
2% Casamino acids (Difco)
0.2% glucose/glycerol (LabServ, BDH respectively)

A.1.3 M9 medium

10.6 g/L Minimal Broth Davis without Dextrose (Difco)
0.2% glucose/glycerol (LabServ, BDH respectively).
100 mg/L thymine

A.2 Media supplements

A.2.1 Amino acids

100 mg/L thiamine
200 mg/L proline

A.2.2 Inducers

1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG; Roche)
L-arabinose (Sigma) (varying concentrations)

A.2.3 Antibiotics

100μg/ml ampicillin
40μg/ml kanamycin
20μg/ml chloramphenicol
15μg/ml tetracycline
100 μg/ml streptomycin

A.3 Solutions for protein extraction and electrophoresis

Protein Lysis Buffer (50 mls)

2.5 ml 1 M Tris EDTA
10 ml 50% glycerol
0.05 ml Triton X-100
37.5 ml ddH20

Laemmli solution (10 ml)

5.67 ml ddH20
2 ml 10% SDS
1 ml glycerol
1 ml 2M Tris p.H 8.4
50 μl 0.1 M EDTA
20 μl 1% bromophenol blue (in EtOH)

Tris-Glycine SDS Buffer 5x (1 L)

15.1 g Tris base
94 g glycine
50 ml SDS (10%)

Appendix B

Sequences

Synthesized Sequences[1]

pBS01 CCC GTG TAA AAC GAC GGC CAG TTT ATC TAG TCA GCT TGA TTC TAG CTG ATC GTG GAC CGG AAG GTG AGC CAG TGA GTT GAT TGC AGT CCA GTT ACG CTG GAG TCT GAG GCT CGT CCT GAA TGA TAT GCG ACC GCC GGA GGG TTG CGT TTG AGA CGG GCG ACA GAT CCA GTC GCG CTG CTC TCG TCG ATC CCT CGA GAA GCT TAT TGA TCG TAT TA

pBB07 TGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTG AGCGGATAACAATTTCACACAGGAAACAGGATCCGTCATAAGAAA GGAGCCGCACATGAAAAAAGCAGTCATTAACGGGGAACAAATCAG AAGTATCAGCGACCTCCACCAGACATTGAAAAAGGAGCTTGCCCT TCCGGAATACTACGGTGAAAACCTGGACGCTTTATGGGATTGTCT GACCGGATGGGTGGAGTACCCGCTCGTTTTGGAATGGAGGCAGTT TGAACAAAGCAAGCAGCTGACTGAAAATGGCGCCGAGAGTGTGCT TCAGGTTTTCCGTGAAGCGAAAGCGGAAGGCTGCGACATCACCAT CATACTTTCTTAATACGATCAATAAGCTTcccgagtttttcagcaagatAGCG GCAAGCTTAATGTGCCTGTCAAATGGACGAAGCAGGGATTCTGCA AACCCTATGCTACTCCGTCAAGCCGTCAATTGTCTGATTCGTTAC CAATTATGACAACTTGACGGCTACATCATTCACTTTTTCTTCACAA CCGGCACGGAACTCGCTCGGGCTGGCCCCGGTGCATTTTTTAAAT ACCCGCGAGAAATAGAGTTGATCGTCAAAACCAACATTGCGACCG ACGGTGGCGATAGGCATCCGGGTGGTGCTCAAAAGCAGCTTCGCC TGGCTGATACGTTGGTCCTCGCGCCAGCTTAAGACGCTAATCCCT

AACTGCTGGCGGAAAAGATGTGACAGACGCGACGGCGACAAGCAA
ACATGCTGTGCGACGCTGGCGATATCAAAATTGCTGTCTGCCAGG
TGATCGCTGATGTACTGACAAGCCTCGCGTACCCGATTATCCATC
GGTGGATGGAGCGACTCGTTAATCGCTTCCATGCGCCGCAGTAAC
AATTGCTCAAGCAGATTTATCGCCAGCAGCTCCGAATAGCGCCCT
TCCCCTTGCCCGGCGTTAATGATTTGCCCAAACAGGTCGCTGAAA
TGCGGCTGGTGCGCTTCATCCGGGCGAAAGAACCCCGTATTGGCA
AATATTGACGGCCAGTTAAGCCATTCATGCCAGTAGGCGCGCGGA
CGAAAGTAAACCCACTGGTGATACCATTCGCGAGCCTCCGGATGA
CGACCGTAGTGATGAATCTCTCCTGGCGGGAACAGCAAAATATCA
CCCGGTCGGCAAACAAATTCTCGTCCCTGATTTTTCACCACCCCCT
GACCGCGAATGGTGAGATTGAGAATATAACCTTTCATTCCCAGCG
GTCGGTCGATAAAAAAATCGAGATAACCGTTGGCCTCAATCGGCG
TTAAACCCGCCACCAGATGGGCATTAAACGAGTATCCCGGCAGCA
GGGGATCATTTTGCGCTTCAGCCATACTTTTCATACTCCCGCCAT
TCAGAGAAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTC
ACTGCGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCCC
GCTTATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGAC
AAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCA
CATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATT
TTTATCCATAAGATTAGCGGATCCTACCTGACGCTTTTTATCGCA
ACTCTCTACTGTTTCTCCATACCCGTTTTTTTGGGCTAGAAATAAT
TTTGTTTAACTTTAAGAAGGAGATATACATACCCATGGAACAAAG
CACTATTGCACTGGCACTCTTACCGTTACTGTTTACCCCTGTGACA
AAAGCCGCACAGGTTATCAACACGTTTGACGGGGTTGCGGATTAT
CTTCAGACATATCATAAGCTACCTGATAATTACATTACAAAATCA
GAAGCACAAGCCCTCGGCTGGGTGGCATCAAAAGGGAACCTTGCA
GACGTCGCTCCGGGGAAAAGCATCGGCGGAGACATCTTCTCAAAC
AGGGAAGGCAAACTCCCGGGCAAAAGCGGACGAACATGGCGTGAA
GCGGATATTAACTATACATCAGGCTTCAGAAATTCAGACCGGATT
CTTTACTCAAGCGACTGGCTGATTTACAAAACAACGGACCATTAT
CAGACCTTTACAAAAATCAGATAACGAAAAAAACGGCTTCCCTGC
GGGAGGCCGTTTTTTTCAGCTTTACATAAAGTGTGTAATAAATTT
TTCTTCAAACTCTGATCGGTCAATTTCACTTTCCGGCTCCAGAGCT

CCAGAGTCCGGTCCAATCTGCAGCCGTCCGAGACAGGAGGACATC
GTCCAGCTGAAACCGGGGCAGAATCCGGCCATTTCTGAAGAGAAA
AATGGTAAACTGATAGAATAAAATCATAAGAAAGGAGCCGCACAT
GAAAAAAGCAGTCATTAACGGGGAACAAATCAGAAGTATCAGCGA
CCTCCACCAGACATTGAAAAAGGAGCTTGCCCTTCCGGAATACTA
CGGTGAAAACCTGGACGCTTTATGGGATTGTCTGACCGGATGGGT
GGAGTACCCGCTCGTTTTGGAATGGAGGCAGTTTGAACAAAGCAA
GCAGCTGACTGAAAATGGCGCCGAGAGTGTGCTTCAGGTTTTCCG
TGAAGCGAAAGCGGAAGGCTGCGACATCACCATCATACTTTCTTA
ATACGATCAATGGGAGATGAACAATATGGAAGTTTAAACCCTGAT
ACAGATTAAATC

RBS:pl__tx__E.coli GTAAAACGACGGCCAGTGAGCTCTTAACTTTAAGA
AGGAGATATACATACCCATGACAACACAACATATTATCGAACCAG
GGCAAGCAGTGCATCAAGCAGCGGCTATTCTTTCTTCTTTGGAGT
ACATCAACCAAGCGGAAGCGCGGAGCCTTGGGCCATTGGCCGAAG
CCGTCGCTAATGCTTTTATGGTGGTGTACTACCAAGCTGAAACGG
GCCGGGCGACACAGGCTGATTTTCAAGAAGCAATGAACGCCTTGC
GCCAAGCGTGCAGCTAAtGACGAAAGGGCGGCCACTCGAACGAGC
TCCCTGTGTGAAATTGTTATCCGCT

PT__XC1 TTAACTTTAAGAAGGAGATATACATACCC**ATGGCAACTTT
GAATCCTACCAACGCAACTCAGGCAGTGCATCATGCGGC
CGTGCAGCTCGCGGCCCTAGACTGGATTGACCAGGATGC
AGCCCGGCAGCTCGGGCCGCTGGCCGAAGCAGTCGCCAA
TGCCTTCATGGTCGTCTTCTACCAGGCCGAGACGGGCCG
AGCGACGCCGGCGGACTTCCGTGAGGCACTGAATGCCGT
GCGCCAGTCGCTTCACCCAGCGTAA**GCCACATACTAGGAAC
AACAAGGGCAGCCAATGGCTGAAGATTTTCAGCCTGATACAGATT
AAATCAGAACGCAG

PT__BP1 TTAACTTTAAGAAGGAGATATACATACCC**ATGGCAACCAT
GCACGATACCGAAGCTCGGGAAGTTATCCATCACACGGC
AATGCAGCTCGCGGCCCTAGAATTCATAGATCCGCACGC
AGCTAAGGATCTATCTGAGATGACCGAAGCCATGGTGAA**

CCTCTTTGTGGTCGTCTTCTATCAGGCCGAAACTGGCCGG
GCGACGCGCTCGGACTTTAGAGAGGCAATGACCGCGGTA
CGCCAAGCGTTGCTGCAGTATGCGGAACAGTTGAAGTGC
CATGGCGGCATTTCTGGGAAGGGGGCTAACGAGGCATGG
TCATGAAAAATCATCCGCACCCCGGCGAATTGCCGCGCGAGGAA
GTTTGTAGCCTGATACAGATTAAATC

PT_AB1 TTAACTTTAAGAAGGAGATATACATACCCA**TGACAACACA**
ACATATTATCGAACCAGGGCAAGCAGTGCATCAAGCAGC
GGCTATTCTTTCTTCTTTGGAGTACATCAACCAAGCGGAA
GCGCGGAGCCTTGGGCCATTGGCCGAAGCCGTCGCTAAT
GCTTTTATGGTGGTGTACTACCAAGCTGAAACGGGCCGG
GCGACACAGGCTGATTTTCAAGAAGCAATGAACGCCTTG
CGCCAAGCGTGCAGCTAATGACGAAAGGGCGGCCACTCGAAC
TGGCCGCCTTTTTCAATCTTCACTGTAAGATTTTCAGCCTGATACA
GATTAAATCAGAACGCAGAA

PR1_EC1 GTAAAACGACGGCCAGTTGGCCAAACAGTGAAGATTGAAA
AAGGCGGCCAGTTCGAGTGGCCGCCCTTTCGTCA**TTAGCTGCAC**
GCTTGGCGCAAGGCGTTCATTGCTTCTTGAAAATCAGCCT
GTGTCGCCCGGCCCGTTTCAGCTTGGTAGTACACCACCAT
AAAAGCATTAGCGACGGCTTCGGCCAATGGCCCAAGGCT
CCGCGCTTCCGCTTGGTTGATGTACTCCAAAGAAGAAAG
AATAGCCGCTGCTTGATGCACTGCTTGCCCTGGTTCGATA
ATATG*TTGTGTTGTCAT*AATTCACCTCTCCAGTAGGTTTTGAT
*TGAAAGTGCCGGGCTGGTTGTGGCCAGCCCGGTGCGCCT*ACAGGT
ACAGCCCCGCCGCCTCCGCGAATGCTGACCGCTCATTCCTGCCGC
CATTAACCTCAATACTTATCGGCTTAGTTTCGCTCTCTCTCATAGG
GTCTGGCCGTGATACAAAAGCGGCTGTTTCTGTCTCTGCATCGAC
CTCGGTCGATTGATCACCCAGCGGCGAAATAAATACCTGGCTGAC
TTCTTCCCTTGGCCCTGTGTGAAATTGTTATCCGCT

PR1_PA1 GTAAAACGACGGCCAGTGAAAACGACAAGGGCGGCCACTG
GCCGCCCTTGCTGTTTCTGGTGGAGC**CTACGCGGTGAGTGAC**
TGCCGCACCGCGTCCAGGGCCTGGCGGAAATCGTCCCGC

GTGGCCTGGCCCGTCTCGGCCTGGTAGTACAGCACCATG
AACATATTGGCGACGGCTTCCGCCATCGGCGAGAGCTGC
CGCGCCGCCTCTTGGTCGATCCAGTCCAGGGCCGCGAGC
TGCGCGGCCGCATGATGGACTGCCTGGGTTGCGATTGCA
TTGGTAGGATTCAAAG*TTGCCAT**GATTCACCTCTGTGGTAGG*
TGGGTTGTGGAAGTGTCGGCCGGCCCCACGCCGGCCGGCGCGCTT
ATCGCTGGCCGCTGCTGTAGATGCCGGCGGCCTCGGCCAGCGGCG
TGTTCCTGTTGCTGCCGCTGTAGCTGTTGCTCGATGACCTGGCCG
CCGGCCTGGTGGCCTCGACGGCAGACGGGGCAGGGCGATCAACGA
AGGCCGCCACTTCCGCCGCTGCATCAAGCTCGGTTGATTGGTCGC
CAGGCTGCTCAAACTCCACCCTGTGTGAAATTGTTATCCGCT

1 PT ORFs indicated in bold, PtaRNA1 antitoxin in italics.

Table B.1: Sequences used in Blastp searches of NCBI database.

Inhibitors	
Barstar	mkkavingeq irsisdlhqt lkkelalpey ygenldalwd cltgwveypl vlewrqfeqs kqltengaes vlqvfreaka egcditiils
Bpi	mkkvqldgal crtqeelhdq lktvlhlpdy ygknldalwd cltgevnlpv eltwvnfdts kdalgeyaes vkqlfkeaee elkgqfqvsi q
Sai	msgleepsad vsavkaaeda errgaaahvl dgsellskra aldgiaavld fpewagrnld alydcltdls wlpegehvli wsgyqalady dpkayrkisa vlkeasetsf cgrtftavlt rn
Sai20	mtdnelivdl rgrqietlnd fwdavsepcg lpewfgrnld awsdtietrg isevidshdi lvvhvdqrgl feghrreadv ladtfdgeqn qlivhgpdwp aatd
Sti	mtvtyvidgf eidtledfwd vvgqaigvdg yfghnldafa dclgggygtp ddgdyviewr hhelsrrrlg cpetvrqlel rlarvhptnr eqvaaelaaa ragkgptvfd wlvdiieerh pg-glrla
RNase	
Barnase	aqvintfdgv adylqtyhkl pdnyitksea qalgwvaskg nladvapgks iggdifsnre gklpgksgrt wreadinyts gfrnsdrily ssdwliyktt dhyqtftkir
Binase	avintfdgv adylirykrl pdnyitksqa salgwvaskg nlaevapgks iggdvfsnre grlpsasgrt wreadinyvs gfrnadrlvy ssdwliyktt dhyatftrir
Bci	aqvintfdgv adylltyhkl pdnyitksea qalgwvaskg nladvapgks iggdifsnre gklpaksgrt wreadinyts gfrnsdrily ssdwliyktt dhyktftkir
Bpu	avintfdgv adylirykrl pdnyitksqa salgwvaskg nlaevapgks iggdvfsnre grlpsasgrt wreadinyvs gfrnadrlvy ssdwliyktt dhyatftrir

Table B.1: Sequences used in Blastp searches of NCBI database continued

Inhibitors	
Sa	dvsgtvclsa lppeatdtln liasdgpfpy sqdgvvfqnr esvlptqsyg yyheytvitp gartrgtrri itgeatqedy ytgdhyatfs lidqtc
Sa3	mripprlval agaaavaatl iagpvaaaap ashavaassa asasvkavgr vcysalpsqa hdtldlideg gpfpysqdgv vfqnregllp ahstgyyhey tvitpgsptr garriitgqq wqedyytadh yasfrrvdfa c
St	laaggsarss ravqalpvap prqapcgdts gfeqvrladl ppeatdtyel iqkggpypyp qdgtvfqnre gilpdcaegy yheytvktpg sddrgarrfv vgdggeyfyt edhyesfrlt ivn

B.1 Type I TA toxin amino acid sequences

Filtered amino acid sequences for families used to derive TMHMM for protein orientation within membranes. Amino acid sequences reported with >Strain accession number/Genomic coordinates Locus identity Bit score E-value.

B.1.1 Hok

>CP001846:5/343201-343243 HOK_GEF 107.1 7.5e-35
KYRLLSLIVICFALLFFTWMIRDSLCELHIKQGSYELAAFLAC
>CP001846:5/1129361-1129403 HOK_GEF 107.1 7.5e-35
KAMLIALIVICLTVIVTALVTRKDLCEVRIRTGQTEVAVFTAY
>CP001846:4/1190772-1190813 HOK_GEF 92.0 3.9e-30
PLVVCLLIICITILTFTLLTRQTLYELRFRDSDKEVAALMAC
>CP001846:4/1790271-1790313 HOK_GEF 92.0 3.9e-30
KVMIVALIVICITAVVAALVTRKDLCEVHIRTGQTEVAVFTAY
>CP001846:0/729193-729234 HOK_GEF 89.5 2.3e-29
QKSLTAITFCVTAILIIWMLHGSLCEIRMSFWGAEFAAFLQC
>CP001846:1/260093-260135 HOK_GEF 88.4 5e-29
KYALVAVIVLCLTVLGFTLLVGDSLCEFTVKERNIEFKAVLAY
>CP001846:3/646993-647035 HOK_GEF 84.0 1.2e-27
KYALVAIIVLCCTVLGFTLMVGDSLCELSIRERGMEFKAVLAY
>CP001891:1/63241-63283 HOK_GEF 58.5 1.1e-19
KYLLFGLVVICFTILLLTWMVRDSLCELQLRQGNIELVAFLAC
>CP001891:0/340039-340081 HOK_GEF 47.3 3.6e-16
KYALVAIIVLCITVLGFTLLVHSSLCELSIKERNIEFKAVLAY
>CP001918:0/289257-289298 HOK_GEF 57.0 3.3e-19
TLLLGLFLICTTLLIFTWMVRDSLCELHFRQEKTELAAVLAY
>CP001918:3/1656530-1656572 HOK_GEF 48.7 1.3e-16

KTALGIVFIICLTIVIFTFITRGKLCELTIKSEHQEVAAKLAC
>CP001918:4/1558646-1558683 HOK_GEF 31.1 4.3e-11
VLMIIALTIIILTLVTRKTLCEIRFRNGSLEVTARMEC
>CP002124:2/962611-962652 HOK_GEF 57.7 2.1e-19
SVIWCLVIVCLTILAFTLLTRHSLYELRIRDGIREVAAVMAC
>CP002272:3/784468-784510 HOK_GEF 49.6 6.6e-17
KAMVVALIVICITVVMAVLVTRKDLCEVRIRIGQMEVTAFTAY
>CP002272:2/330953-330995 HOK_GEF 48.1 2e-16
KYAVVAIIALCITALGFTLMVRDSLCELSIRERGMEFKAVLAY
>CP002729:0/404634-404674 HOK_GEF 37.7 3.5e-13
KTIIVGMLCLTMLLTVWVLHASPCEFRVSFMWSEIAAFLQC
>CP002773:4/323495-323535 HOK_GEF 124.9 2e-40
SVVKLVVICATVISLVWITRSKICELRIRSGNTEVAAILAY
>CP002773:4/582870-582912 HOK_GEF 124.9 2e-40
KHVFHSLTVVCITILLFIWMIRDSLCELAIYQENITILIRLAC
>CP002773:4/1490810-1490847 HOK_GEF 124.9 2e-40
RLMIICMTLIPFMLITRGSLCELRIKLGDTEVAAILAY
>CP002773:3/1490607-1490647 HOK_GEF 48.0 2.2e-16
VVLRLMIICMTLIALMLITRGSLCELHIKLGDTEVAAILAY
>CP002797:0/1629814-1629855 HOK_GEF 54.0 3e-18
ALIWCVLIVCCTLLIFTLLTRNRLCEVRLKDGYREVTATMAY
>CP003218:0/998205-998247 HOK_GEF 47.4 3.4e-16
KYALVAIIVLCLTALGLTLMVRDSLCELSIKERSMEFKAVLAY
>CP003218:4/1919692-1919734 HOK_GEF 44.1 3.7e-15
KYALVAIVVLCITVLGFTLLVRSSLCELSTKERSMEFKAVLAY
>CU928162:5/1433533-1433572 HOK_GEF 93.9 9.7e-31
TIIVSILCLTMLLAIWMIHRSPCEFRLNIMWSEFAAFLQC
>CU928162:2/382658-382699 HOK_GEF 90.3 1.4e-29
QKSLITVTICMTVIFTIWMLHGSLCEFRLNLWGAEFAAFLQC
>FM162591:3/463959-464001 HOK_GEF 103.5 1e-33
KIAIFSLIIICVTLLIFTWITRGSLCKLHIKRGNTEVAAIMAY
>FM162591:3/974182-974224 HOK_GEF 103.5 1e-33
KIVIFSLIVICVTLLLFTWITRGSLCKLHFKQGNTEVAAIMAY
>FM178379:3/1101138-1101180 HOK_GEF 58.9 8.3e-20
KSALTGLIVICLTILCFTWMVRGSLCELQIKNGSAVIQATLAY
>AM942759:0/633875-633917 HOK_GEF 76.6 2.5e-25
KLALIGLITVCITVLCFTLLRHERLCSFNISSGNTLVQANLSC
>AM942759:1/12344-12386 HOK_GEF 40.8 3.7e-14
KITLFGITVVCVTVLMFALIVGERLCSLNISNGNMVVQAILMC
>FN434113:3/205121-205162 HOK_GEF 65.5 7.6e-22
CLIWCVLIVCVTLLIFTYLTRKSLCEIRYKDGYREVAAFMAY

>FN543093:3/1127904-1127945 HOK_GEF 99.1 2.3e-32
AFIWCIIIVCVTVLVFTTLSRETLCELRLRGASMEIVASLAC
>FN543093:3/1447455-1447497 HOK_GEF 99.1 2.3e-32
STLFWCVLMICITLLMFTFLTRHSLCELRLKDGQREFAAFLAY
>FN543093:5/53045-53086 HOK_GEF 93.9 1e-30
YAFLSIVALCITALCFALIMRDRLCEFTIKEGSREIRAILAY
>FN543093:5/53432-53472 HOK_GEF 93.9 1e-30
LIWCVLIVCITLLIFTLLIRKSLCEIRYKDRYREAAAFLAY
>FN543093:0/1295352-1295393 HOK_GEF 53.9 3.2e-18
YLLACLFMVCITILIFALMNQGTLCELTIRSGSQEVAAKLAC
>FN543093:2/1295525-1295567 HOK_GEF 52.1 1.2e-17
KYLFACFVVFCVTILIFALINHGTLCELTIKQGNKEVAARLSC
>FN543502:1/1493949-1493991 HOK_GEF 63.9 2.3e-21
KCILCSLIVICLTILLFTWMVRDSLCELHIRQGNNELAAFLAC
>FN543502:0/213098-213140 HOK_GEF 47.9 2.4e-16
KYALMAIIVLCVTVLGFTLLVGDSLCEFSVKERGMEFKAVLAY
>FP929040:0/1214927-1214969 HOK_GEF 62.9 4.8e-21
NTLLWCVLIVCLTLLIFTYLTRKSLCEIRYRDTDREVAAFMAY
>FR729477:3/812458-812500 HOK_GEF 57.4 2.5e-19
NTLVLCLLIVCITVLAFTLITHKSLCELRVKNGIKEVAAILAY
>AB255436:0/8420-8459 HOK_GEF 38.8 1.6e-13
TIIVSILCITILVIVWMIHNSHCEFRLNIMWSQFAAFLQC
>BX950851:1/944119-944158 HOK_GEF 42.1 1.5e-14
FFSLIVLCLTLLVFTSMVRGSLCKIRVQQGNTVMVASLNY
>CP000826:1/1335009-1335047 HOK_GEF 40.5 4.7e-14
PGFTVGCVTFLVFIWLIRGSVCELTFKQGNTEVAAVLSC

B.1.2 FlmA

>AE005174.5/331007-331060
NQHRGLGMPQKYRLLSLIVICFTLLFFTWMIRDSLCELHIKQGSYELAAFLACN
>AP009378.4/346499-346552
KSTQEAGMPQKYRLLSLIVICFTLLFFTWMIRDSLCELHIKQENYELAAFLACK
>CP000647.3/338612-338660
MPQKYLLFGLVVICFTILLLTWMVRDSLCELQRRQGNIELVAFLACDIK
>CP003218.5/1603101-1603149
RRGMPQKYLLFSLIVICFTILLFTWMVRDSLCELQLRQGNIELVAFLAC
>CP001918.0/289221-289301
RVPHQKTDGETNFTTNIARLVSSLPKSNKEAGMPKRTLLLGLFLICTTLLIFTWMVRDSLCELHFRQEKTELAAVLAYEAK
>AY258503.3/40466-40523
RPKGNGGAAMPQRTFLTMLIVVCVTILCFVWMVRDSLCGFRVEQGNTVLVATLAYEVK
>EU935740.1/15295-15351

PEGNGGAMMPQRTFLMMLIVVCVTILCFVWMVRDSLCGLRLQQGNTVLVATLAYEVK
>CP002775.3/582854-582905
KEASMRLKHVFHSLTVVCITILLFIWMIRDSLCELAIYQENITILIRLACEV
>AE005174.0/629533-629601
LNTCRLASYVPKGKEKQAMKQQKAMLIALIVICLTVIVTALVTRKDLCEVRIRTGQTEVAVFVDYESEK
>CU928162.4/1167350-1167417
LVTCRLASYTPKGKEKQAMKQQKAMLVALIVICITVIVTALVTRKDLCEVRIRTGQTEVAVFTAYEPE
>AE005174.5/917035-917101
HKRRLASYAPKGKEKQVMKQQKAMLIALIVICLTVIVTALVTRKDLCEVRIRTGQTEVAVFTAYEPE
>AE005174.5/1837597-1837674
FYHQPEAPYMLNTCRVPLTDRKVKEKRAMKQHKVMIVALIVXCITAVVAALVTRKDLCEVHIRTGQTEVAVFTAYESE
>CP002775.4/1490515-1490603
TDRKRKAPSRYYQPEATLNVQTTSGPLTRRKARIRRSEMQQERVVLRLMIICMTLIALMLITRGSLCELHIKLGDTEVAAILAYESER
>CP002775.5/1490754-1490802
MQQERVIPRLMIICMTLIPFMLITRGSLCELRIKLGDTEVAAILAYESE
>CP002775.3/323480-323536
RQGEGGMMPDKRSVVKLVVICATVISLVWITRSKICELRIRSGNTEVAAILAYESKR
>CP000799.1/24230-24271
IGLLAVCATVLCFSLIFRERLCELNIHRGNTVVQVTLAYEAR
>AP009240.4/1103534-1103583
KAMKHNPLVVCLLIICITILTFTLLTRQTLYELRFRDGDKEVAALMACTS
>CP002124.2/962565-962656
FQAYTKQKEKSPELINQPEALVCSTHARLASYRPKGQGEAAMKPQSSVIWCLVIVCLTILAFTLLTRHSLYELRIRDGIREVAAVMACESSQ
>AY214164.3/19446-19525
FINHHEASLCLVHIRIASYRALRKEKKAMKLPRSSLVWCVLIVCLTLLIFTYLTRKSLCEIRYRDGHREVAAFMAYESGK
>CP001369.2/16960-17048
QKDKSPVVNFSLTHEAPACLDNIRIASYRALRKEKKAMKLPRSSLVWCVLIVCLTLLIFTYLTRKSLCEIRYRDGYREVAAFMAYESGK
>FJ876826.1/14048-14107
PQGKEKKAMKLPRNTLLGCVLIVCLTLLIFTYLTRKSLCEIRYRDTNREVAAFLAYESAK
>CP001122.3/30636-30687
TMKLPGNALIWCVLIVCCTLLIFTLLTRNRLCEVRLKDGYREVTASLAYESS
>FN434113.3/205092-205166
EAICMPYTSRIASYPPQGKDEKAMKLPANCLIWCVLIVCVTLLIFTYLTRKSLCEIRYKDGYREVAAFMAYESGK
>AB366440.4/18402-18470
VTFRVASYPLKCKEKKAMKLPNQPVVLCVLIVCLTLLIFTWLTRNSLCELRLKDGTREVSAVMDYESGK
>EF382672.5/25774-25828
EKKAMKMPNQPIVLCILIVCLTLLIFTWLTRNSLCELRMKDGTREVSAILAYESG
>CP001918.4/1558634-1558687
EASMKRQSPIIRVLMIIALTIIILTLVTRKTLCEIRFRNGSLEVTARMECRSGQ
>CP000893.1/16165-16209
RKTTTLSLAIVCFTLLAALALIRDDLCKVEYRNGAMLLNVVLAYE

>CP001918.3/1656523-1656573
VKGMTPLKTALGIVFIICLTIVIFTFITRGKLCELTIKSEHQEVAAKLACV
>AE016796.0/238569-238615
MPKKTALAALVVVCITAVILTALYTGSLCDIRYKDQQNDLSIKLAYE
>AE005174.5/617912-617954
VAIIVLCCTVLGFTLMVGDSLCELSIRERGMEFKAVLAYESKK
>CP000647.3/640242-640284
VAIIVLCITVLGFTLLVHSSLCELSIKERNIEFKAVLAYESKK
>CP003218.4/1919696-1919738
VAIVVLCITVLGFTLLVRSSLCELSTKERSMEFKAVLAYESKK
>CP003218.0/998205-998251
KYALVAIIVLCLTALGLTLMVRDSLCELSIKERSMEFKAVLAYETKK
>CP001846.1/260095-260139
ALVAVIVLCLTVLGFTLLVGDSLCEFTVKERNIEFKAVLAYEPKK
>AB255436.0/8416-8460
MPQKTIIVSILCITILVIVWMIHNSHCEFRLNIMWSQFAAFLQCK

B.1.3 Fst

>CP001993:4/527346-527366 Fst_toxin 23.6 1.1e-08
LKTIIGPIVVGVVLRIVDKWL
>CP002340:0/308094-308114 Fst_toxin 27.8 5.1e-10
FKTIIGPIVVSIILRLVDKWL
>CP002429:1/701095-701115 Fst_toxin 25.0 4.1e-09
FSSIVAPIIVGIVLALFEHWL
>CP002616:3/47132-47152 Fst_toxin 21.7 4.2e-08
LNQIIAPLIVGVVLLLLEHRL
>CP002888:2/368655-368675 Fst_toxin 31.8 2.9e-11
FKTIIGPIVVGVILRLIDKWL
>FN298497:0/363182-363202 Fst_toxin 11.9 5.1e-05
LTLVVAPILVEIVKSLFDHWL
>FN557490:3/695194-695214 Fst_toxin 22.9 1.9e-08
LSLVIAPIFVGIVIQLVSHWL
>FR821777:0/352621-352641 Fst_toxin 16.1 2.5e-06
LQLIIGALVVGMICLAVSHWF
>AP008934:5/539806-539826 Fst_toxin 26.4 1.4e-09
FVTFIAPIVVGVIVTLFSYWL
>DQ320509:4/76-96 Fst_toxin 28.2 3.8e-10
LTSVVAPILVGIVLALFDHWL
>CP000017:1/392003-392023 Fst_toxin 24.9 4.3e-09
FTTIIAPLLVGIILLLIQKWL
>AB370336:5/90-104 Fst_toxin 16.3 2.1e-06

IIAPLLVGVILLLVE
>CP000423:0/889693-889713 Fst_toxin 29.6 1.4e-10
FFQIIAPLLVGVMVTLVDHWL
>CP000423:5/774897-774917 Fst_toxin 22.6 2.2e-08
FTVVVGPVVVGVIIELIKRWL
>CP000517:2/497039-497058 Fst_toxin 18.9 3.2e-07
SLIVAPIIVGIVLALFNHWL

B.1.4 Ldr

>CP001846:3/351846-351880 Ldr_toxin 125.3 2.1e-40
MTLAELGMAFWHDLAAPVIAGILASMIVNWLNKRK
>CP001846:3/1284415-1284449 Ldr_toxin 125.3 2.1e-40
MTLAQFAMTFWHDLAAPILAGIITAAIVGWWRNRK
>CP002272:5/1542074-1542107 Ldr_toxin 51.3 2.7e-17
MTLAQFGIAFWHDIAAPTVAAIIASLIVGWLESR
>CP002487:4/349467-349501 Ldr_toxin 68.1 1.6e-22
MTLTQLGVVFWHDLAAPIIAGIIASVIVNWLRDRK
>CP002487:1/155611-155645 Ldr_toxin 61.1 2.3e-20
MTLTELSITIWHDLAAPTLVGIATGLFLGWWHRRK
>CP002797:4/1195252-1195286 Ldr_toxin 61.1 2.4e-20
MTLAQFAMIFWHNLAAPILAGIITAVIVSWWRNRK
>CP002824:5/1348311-1348345 Ldr_toxin 48.0 2.9e-16
MTLTQLGLILWDDLAAPIIAGIVVSIIVSWMHNQK
>CP003218:4/1616137-1616171 Ldr_toxin 75.0 1.1e-24
MTLAQLGLLLWDDLAAPVIAGILVSIIVSRMDNRK
>FN543502:1/1501980-1502014 Ldr_toxin 66.7 4.4e-22
MTLIQLGIAFWHDLAAPAIAGIIAGMVVNWLRSRK

B.1.5 ShoB

>CP001846:4/727742-727767 ShoB 52.6 7.9e-18
%MTDCRYLIKRVIKIIIAVLQLILLFL

B.1.6 Ibs

>CP001846:4/589293-589311 Ibs_toxin 62.7 6.2e-21
MMRLVIILIVLLLISFGAY
>CP001846:4/589424-589442 Ibs_toxin 62.7 6.2e-21
MMRLVIILIVLLLISLPAY
>CP001846:4/925792-925810 Ibs_toxin 62.7 6.2e-21
MMKHVIILVILLVISFQAY
>CP001846:1/1274808-1274826 Ibs_toxin 28.7 3e-10

MMKLVIILVVLLLLSFPTY
>CP001846:0/1274683-1274701 Ibs_toxin 22.3 3e-08
MMKLVIILIVLLLVSFAAY
>CP001846:3/925683-925700 Ibs_toxin 15.6 4.1e-06
MMKILIIVVLLVISYPAY
>CP001855:5/857243-857260 Ibs_toxin 23.0 1.9e-08
MMKWIIIVVLLVISFPAY
>CP001969:5/913141-913158 Ibs_toxin 42.3 1.6e-14
MMKLLIIVVLLVISWPAY
>CP002185:3/549092-549110 Ibs_toxin 16.2 2.6e-06
MMRLAIILIVLLLISFPAY
>CP002487:0/1127682-1127700 Ibs_toxin 19.1 3.2e-07
MMHQVIILIVLLLISFAAY
>CP002487:4/888766-888784 Ibs_toxin 18.3 5.7e-07
MMKVVIILVILLVISLPAY
>CP002824:3/88352-88370 Ibs_toxin 23.4 1.4e-08
MMNLLIILIVLLLISFPTY
>CU928158:1/1021444-1021462 Ibs_toxin 25.1 4.2e-09
MMKFVIILIVLLLLSFPTY
>CU928158:0/1021319-1021337 Ibs_toxin 24.8 5.2e-09
MMKLVIILIVLLLISFHAY
>CU928158:4/797232-797249 Ibs_toxin 20.9 8.3e-08
MMKWIIIVVLLVVSYPAY
>FN543502:3/1012446-1012464 Ibs_toxin 25.6 2.9e-09
MMKLVIILVIFLVISFPAY
>FR877557:5/781965-781983 Ibs_toxin 21.1 7.4e-08
MMKVVIILVIRLVISFQAY
>CP000822:2/229108-229126 Ibs_toxin 27.9 5.4e-10
MMKCVIILVILLVISFQTY
>CP000880:3/84055-84073 Ibs_toxin 21.5 5.4e-08
MMQQVIILIVLLQISFPAY
>CP000946:2/285265-285283 Ibs_toxin 50.1 5.8e-17
MMRFVIILIVLLLISFPAY

B.1.7 TisB

>CP001846:1/1510508-1510535 TisB_toxin 39.2 1.6e-13
MSLVDIAILILKLIVAALQLLDAVLKYL
>CP001891:4/1809686-1809713 TisB_toxin 49.0 1.3e-16
MGGMDIIILILKLMVAVLQLLDAVLKQF
>CP001918:3/1761479-1761506 TisB_toxin 39.9 9.3e-14
MSVVDMVILILKLIVAVLQLLDAVLKYL

>CP003218:1/446934-446961 TisB_toxin 44.5 3.2e-15
MGGVDIVVLILKLMVAVLQLLDAVLKQF
>FN543093:3/1441267-1441294 TisB_toxin 38.7 2.1e-13
MSGVDLFILILKLIVAVLQLLDAVLKFL
>FN543502:3/349300-349327 TisB_toxin 40.8 4.8e-14
MSLVDIIILILKLIVAVMQLLDAVLKYF
>CP000653:3/1497378-1497405 TisB_toxin 34.2 5.5e-12
MGAVDMVILILRLIVAILQLLDAVLKYL
>AE017220:1/1312700-1312727 TisB_toxin 43.8 5.5e-15
MRVVDITILILKLIVAALQLLDAVLKYL

B.1.8 Plasmid_toxin

>E2PKN4_9RHIZ/1-74 E2PKN4.1
MATINPKQFSHAVQQAAAQLNAIDWLDQDAARELGPLAEATANLFMVLFYQAETGLATREDFSLARTAIQLALH
>A7MAH3_PSEAE/1-76 A7MAH3.1
MATLNPTNAIATQAVHHAAAQLAALDWIDQEAARQLSPMAEAVANMFMVLYYQAETGQATRDDFRQALDAVRQSLT
>A1W8W7_ACISJ/1-74 A1W8W7.1
MATLNPTNATQAVHHAAVQLAALDWLDQDAARQLGPLAEAVANAFMVVFYQAETGQATPADFREALDAVRQSLG
>Q3AT57_CHLCH/11-71 Q3AT57.1
SLHIKMQEVARQNGVSLDQFIATAIAEKLAALMTVNYLRERTERSSQEDFERALSEIPDVA
>E7I0C2_ECOLX/1-74 E7I0C2.1
MTTTHNIEVQQAIHQAAIKLAALEFIDQETARQVLPVAEAVANMFTILYYQAETGRVTPEDFEQALDTIRQVTK
>A1U1D1_MARAV/1-74 A1U1D1.1
MTATHNIEVQQAIHQAATKLSALKFIDQETARQISPVAEAVANMFTILYYQAETGRATQEDFQEALATIRQATE
>A0DJC3_PARTE/8-46 A0DJC3.1
YEDEFQKYSKALMSKFHQAPTATQQDYREAMEEIQQFEK
>E7JDT7_ECOLX/1-74 E7JDT7.1
MTTQHIIEPGQAVHQAAAILSSLEYINQAEARSLGPLAEAVANAFMVVYYQAETGRATQADFQEAMNALRQACS
>B4EEA0_BURCJ/1-74 B4EEA0.1
MATMHDTEAREVIHHTAMRLAALEFIDRHAAKDLSQMAEAVANLFVVVFYQAETGRATRSDFREAMTAVRQALQ
>C0YE05_BURPS/1-49 C0YE05.1
MDQRTARELSKMAEAVANLFMVVFYQAETGRATRLDFSEAMAAVRETLQ
>D2UDH0_XANAP/1-74 D2UDH0.1
MATTHDIEARQVIHHAATQLADLDFMDQRTAQELSTMAEAVANLFMVVFYQAETGRATHGDFSEAMAAVRQTLK

B.1.9 TxpA

>AL009126:1/892747-892805 TxpA 137.4 4.3e-44
%MSTYESLMVMIGFANLIGGIMTWVISLLTLLFMLRKKDTHPIYITVKEKCLHEDPPIKG

Appendix C

Supplementary Data

C.1 Chapter 3

C.1.1 Sequence logos for type I TA toxin families and aggregated HMMs

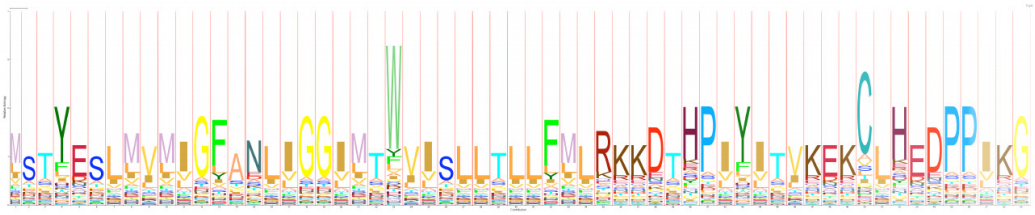


Figure C.1: TxpA

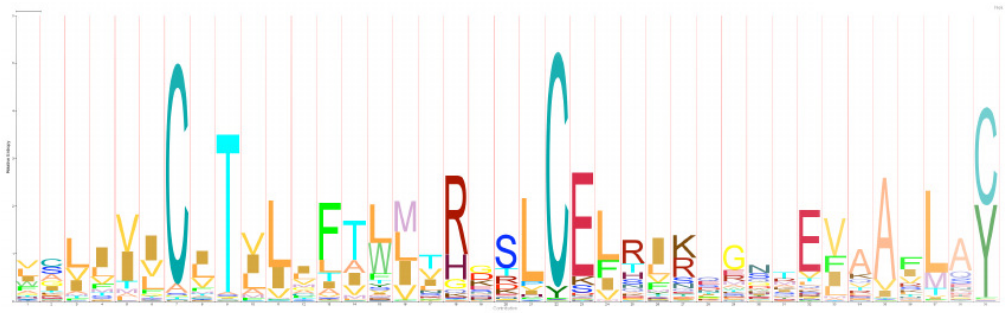


Figure C.2: Plasmid_Toxin

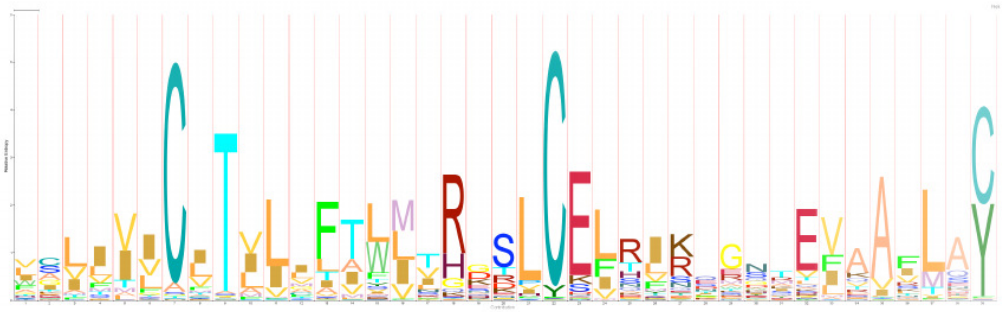


Figure C.3: Hok

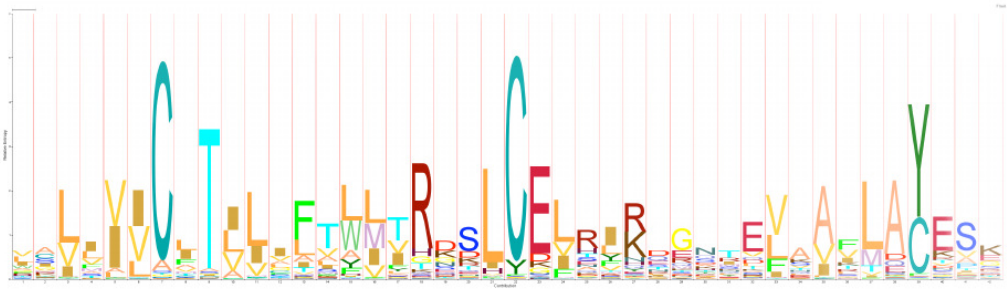


Figure C.4: FlmA

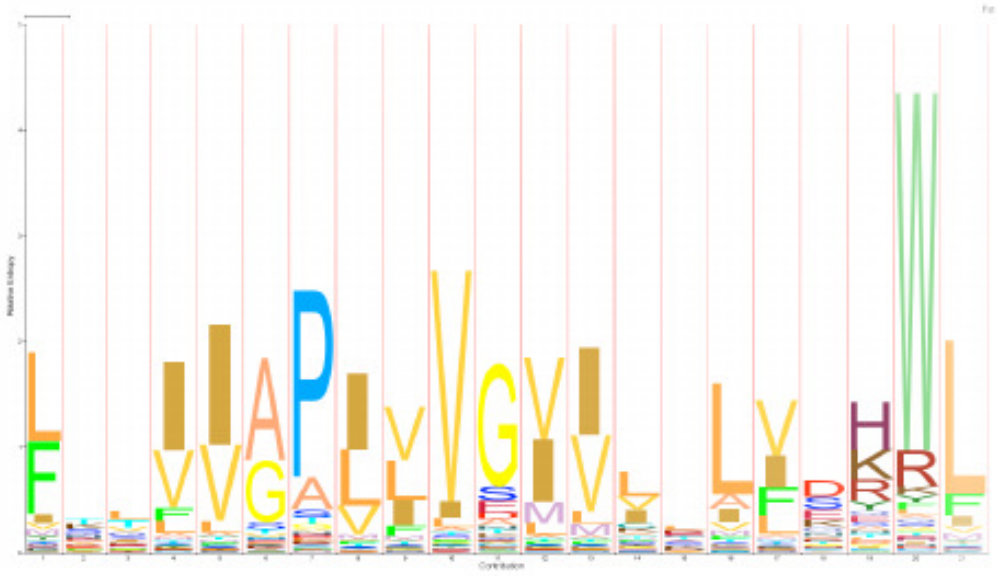


Figure C.5: Fst

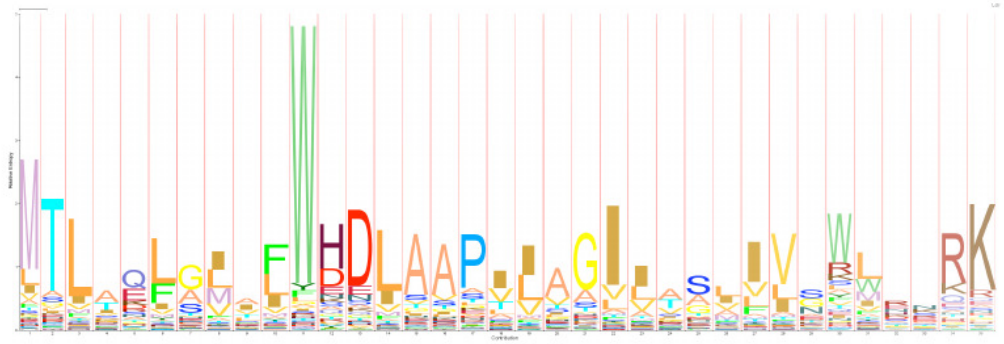


Figure C.6: Ldr

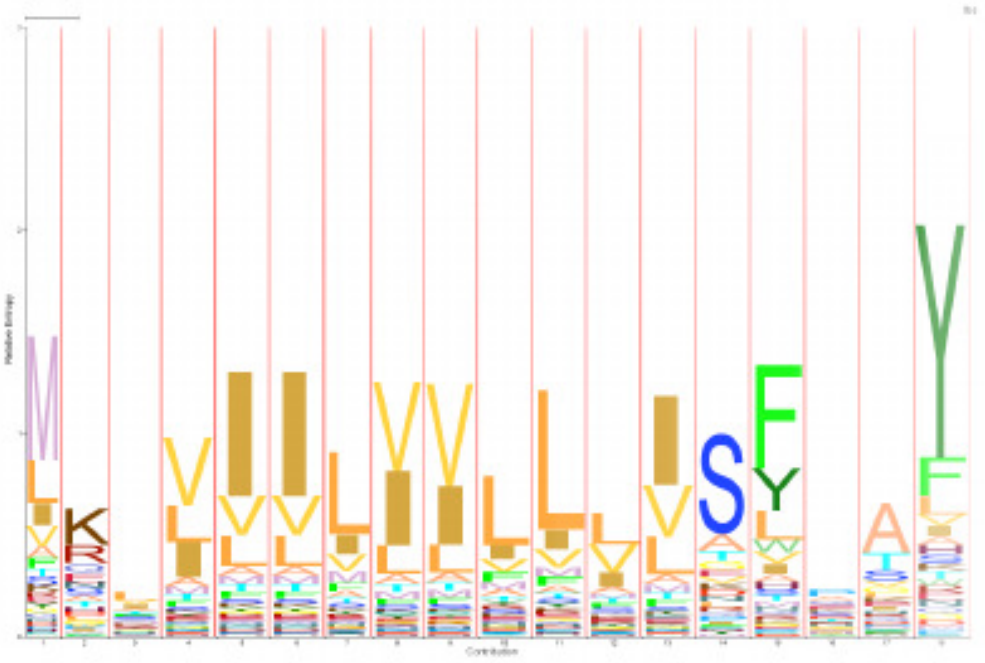


Figure C.7: Ibs

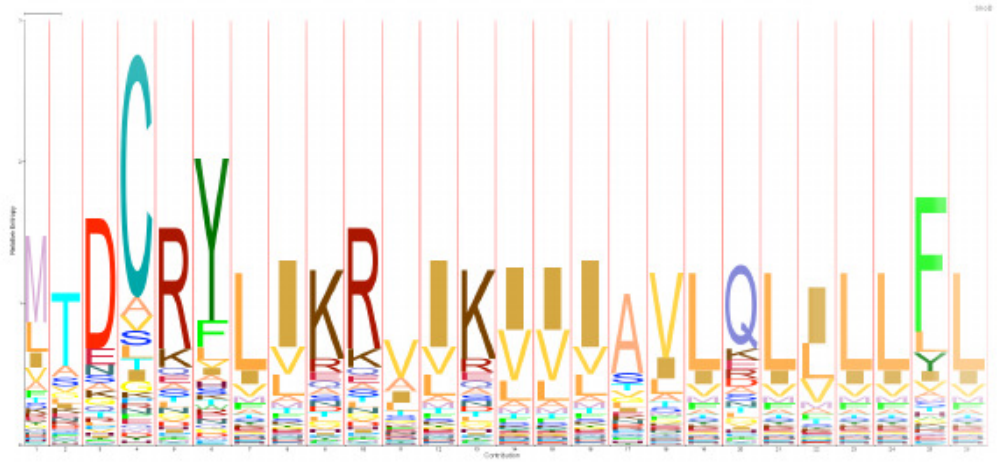


Figure C.8: ShoB

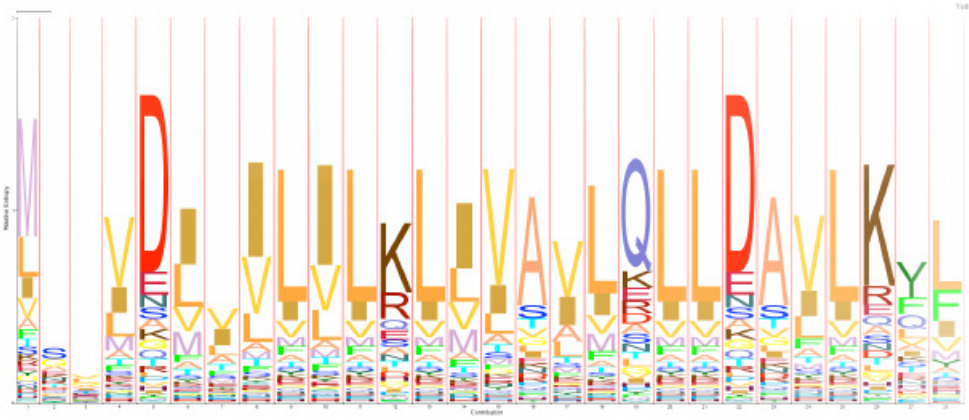


Figure C.9: TisB

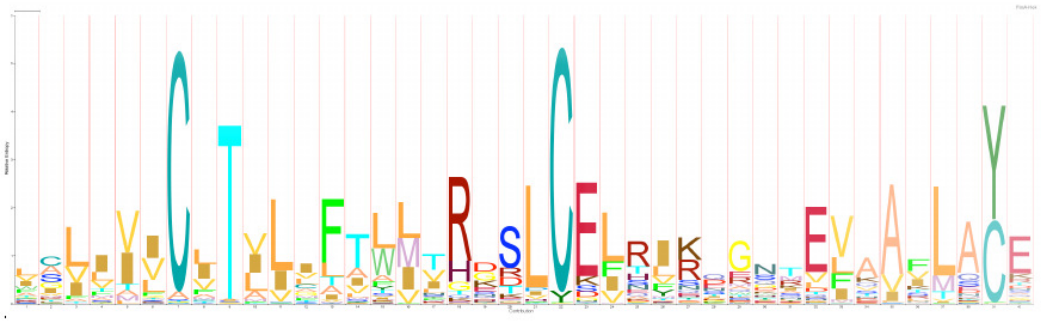


Figure C.10: Aggregate of Hok and FlmA families

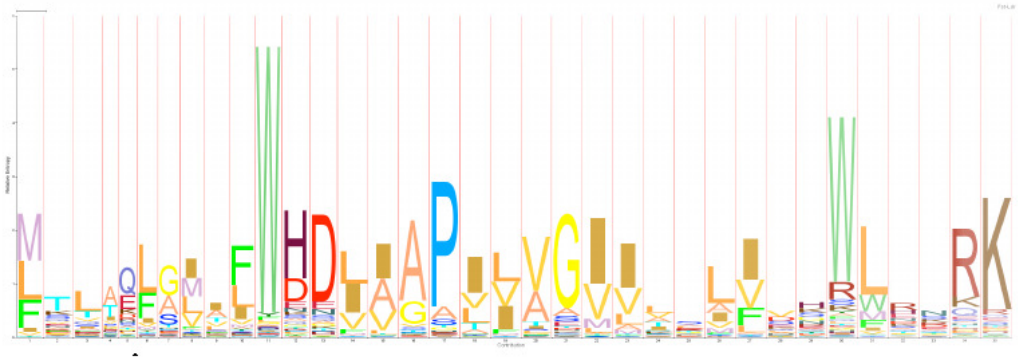


Figure C.11: Aggregate of Fst and Ldr families

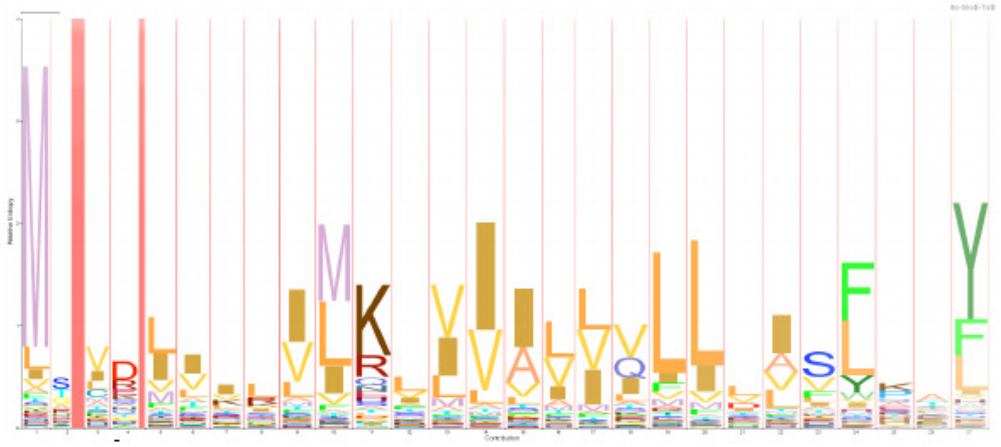


Figure C.12: Aggregate of Ibs, TisB and ShoB

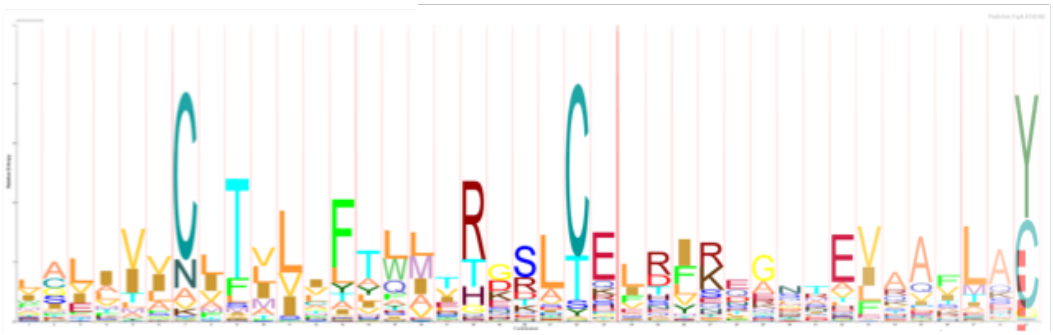


Figure C.13: Aggregate of TxpA, plasmid_Toxin, Hok and FlmA families

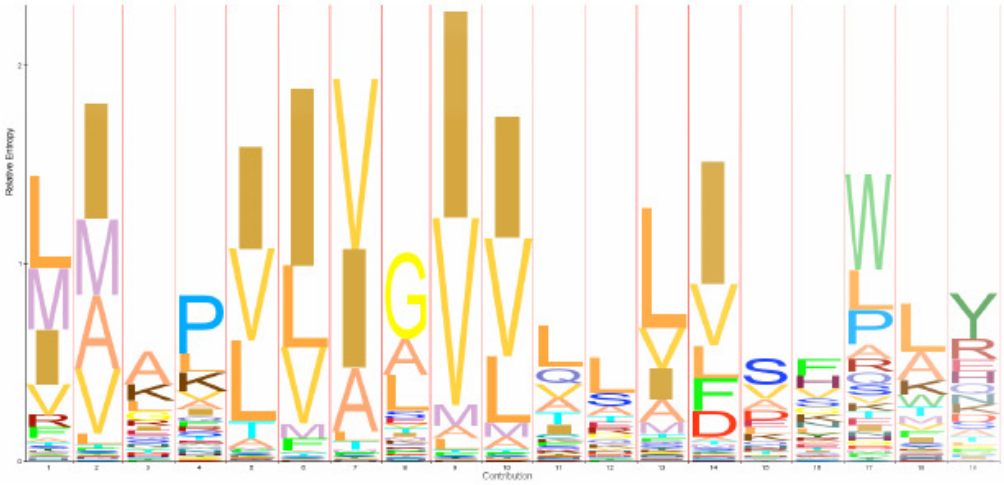


Figure C.14: Aggregate of Fst, Ldr, Ibs, TisB and ShoB

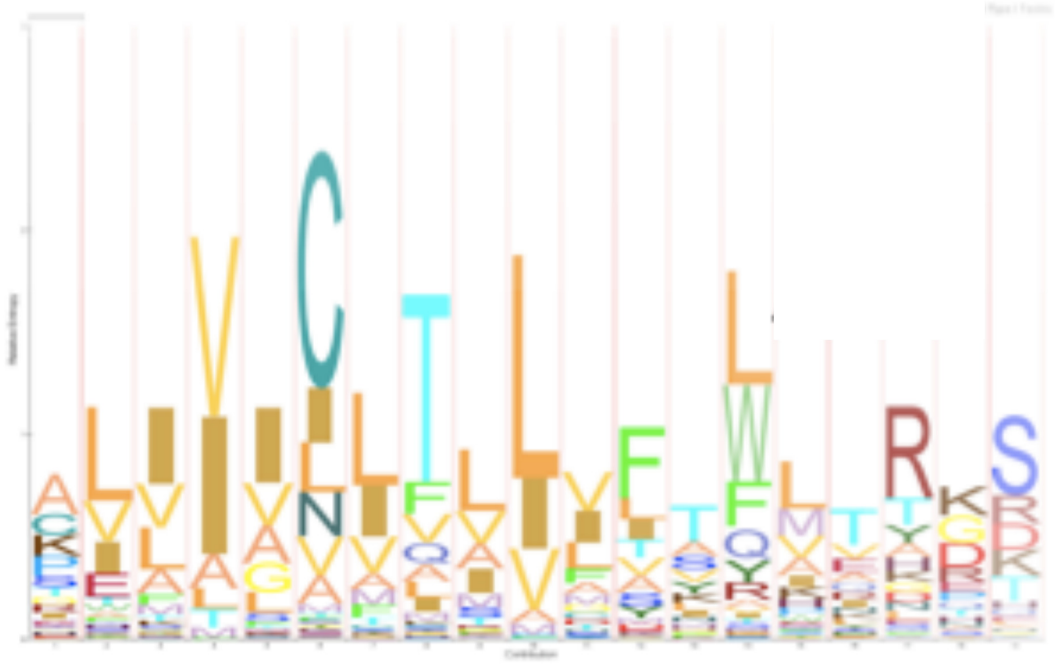


Figure C.15: Aggregate of all families

C.2 Chapter 5

C.2.1 Protein profiles of CSP01 and CSBs01

Cultures of CSP01 and CSBB01 were grown to mid-log phase before treatment with IPTG to induce expression of the *lacZ* fragment and barnase, respectively. Cultures were sampled at 10, 30, and 60-minute intervals for three hours (Figures B.1 and B.2). Proteins from total cell extracts were separated with SDS-PAGE, using a 4-20% gradient gel. Samples were normalized by running equal volumes on a gel, and calibrating protein content using the computer program GeneTools.

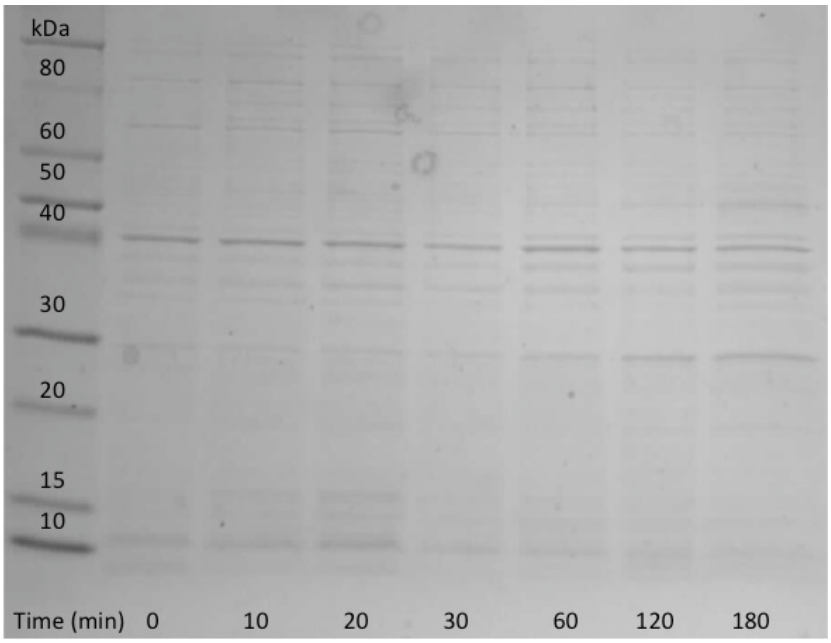


Figure C.16: Intracellular protein profile of cells expressing a *lacZ* fragment. *E. coli* strain CSP01 were grown to ~0.500 $OD_(600)$ before induction (1.0 mM IPTG) and sampling. Proteins were extracted from cell fractions, and content was normalized between samples.

In the CSP01 controls, the induced *lacZ* fragment is 52 amino acids long and does not appear on the gel. The ~30 kDa protein that appears to accumulate in CSP01 cells could be β-lactamase, but would be expected in the other gels as well, as all cell types were grown in ampicillin.

Protein extracts of uninduced CSBB01 were also separated with SDS-PAGE. Barnase accumulated in these cells, producing a band at ~12.5 kDA, though not to the same degree as in induced cells (Figure B.2).

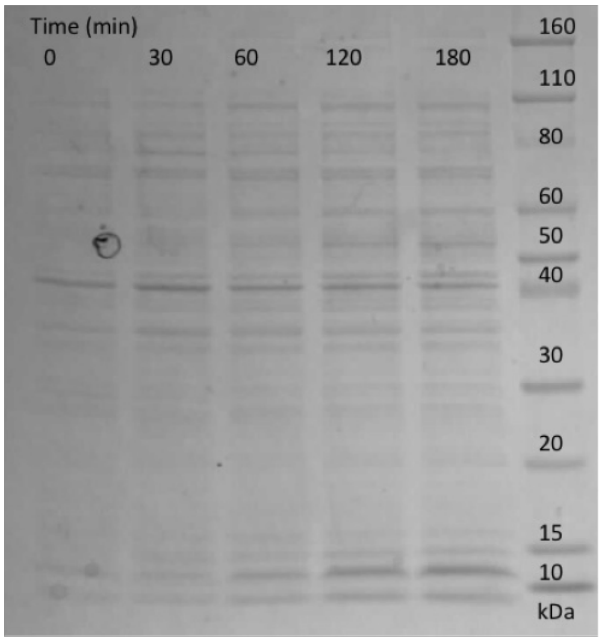


Figure C.17: Intracellular protein profile of barnase-containing cells, uninduced. *E. coli* strain CSBB01 was grown to ~0.500 $OD_{(}600$) before sampling. Proteins were extracted from cell fractions, and content was normalized between samples.

C.3 Chapter 6

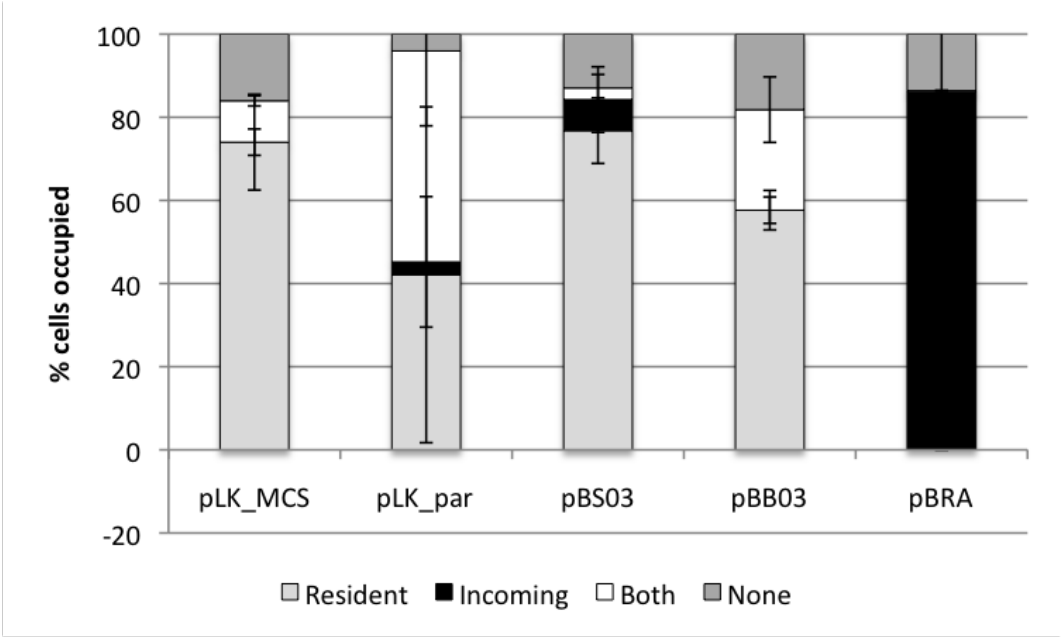
C.3.1 Testing PSK by secreted barnase by culturing double-transformed E.coli

Stability in monoculture is a test of PSK. Cells containing either a resident +PSK (pLK_par), -PSK (pLK_MCS, pBRA) or test (pBB03, pBs03) plasmid were transformed with a second incoming plasmid, pBRT to see if the incompatible plasmid increased the rate of plasmid loss. Cultures were grown overnight in antibiotics to select for a homogenous population containing both plasmids, then cultured for 50 generations without selection. We would expect that a culture with a –PSK resident plasmid would, as the plasmids segregated, contain roughly equal numbers of the resident and incoming plasmid. A +PSK resident culture should loss all segregants with the incoming plasmid.

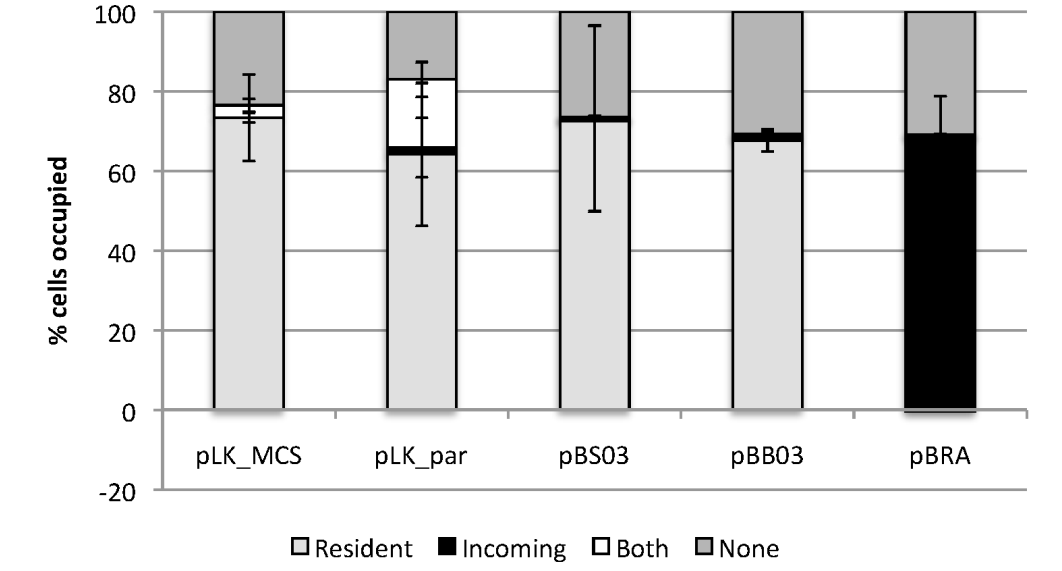
The number of cells containing both did decrease as the generations increased, from 100% at the initial time point (data not shown), suggesting segregation of the plasmids (Figure B.3). Interestingly, all cells with resident plasmids containing kanamycin resistance markers (all but pBRA) primarily retained the resident plasmids. The negative control with a resident ampicillin resistant plasmid primarily retained the incoming plasmid.

The high level of deviation in some conditions was the result of different plasmid dynamics in different independent replicas- some cultures of pLK_par, for example, had mostly resident plasmid at 20 generations, the pattern that was eventually seen in all cultures by 50 generations (Figure B.3).

Previous work in this lab has shown fluctuating plasmid dynamics when incompatible, multicopy plasmids are in the same cell, sensitive to the selection pressures used to measure the culture (Chapter 5, and Ryan Catchpole, personal communication). This suggests that a lower copy number system would be more effective for such a test. Also, plasmid loss can be increased by using greater volumes and larger dilutions, up to 10^6 dilution instead of the 10^3 used here. This increases the amount of time cells are in exponential phase, which tends to increase the rate of plasmid segregation.



(a)



(b)

Figure C.18: Percentage of cells retaining +/- PSK plasmids during serial culture without selection. *E. coli* strains containing +PSK (pLK_par), -PSK (pLK_MCS, pBRA) and two test plasmids (pBS03, pBB03) were transformed with an incompatible plasmid. Resulting double transformants were serially cultured without selection and tested for plasmid composition. A) After 20 generations B) After 50 generations. Results are of one trial with four independent replicas of each strain within trial.

Bibliography

Afif, Hassan, Noureddine Allali, Martine Couturier, and Laurence Van Melderen (2001). "The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system". In: *Molecular Microbiology* 41.1, pp. 73–82. ISSN: 1365-2958.

Aizenman, Einat, Hanna Engelberg-Kulka, and Gad Glaser (1996). "An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3', 5'-bispyrophosphate: a model for programmed bacterial cell death". In: *Proceedings of the National Academy of Sciences* 93.12, pp. 6059–6063.

Alexeyev, M. F. and I. N. Shokolenko (1995). "Mini-Tn10 transposon derivatives for insertion mutagenesis and gene delivery into the chromosome of Gram-negative bacteria". In: *Gene* 160.1, pp. 59–62.

Amábile-Cuevas, Carlos F. and Jack A. Heinemann (2004). "Shooting the messenger of antibiotic resistance: plasmid elimination as a potential counter-evolutionary tactic". In: *Drug discovery today* 9.11, pp. 465–467. ISSN: 1359-6446.

Ambur, Ole Herman, Stephan A. Frye, Mariann Nilsen, Eirik Hovland, and Tone Tønjum (2012). "Restriction and sequence alterations affect DNA uptake sequence-dependent transformation in *Neisseria meningitidis*". In: *PloS one* 7.7, e39742. ISSN: 1932-6203.

Amitai, S., Y. Yassin, and H. Engelberg-Kulka (2004). "MazF-mediated cell death in *Escherichia coli*: a point of no return". In: *Journal of Bacteriology* 186.24, p. 8295.

Aravind, L and Eugene V Koonin (2001). "The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate-and iron-dependent dioxygenases". In: *Genome Biol* 2.3, pp. 1–0007.

Baquero, Fernando, José-Luis Martínez, and Rafael Cantón (2008). "Antibiotics and antibiotic resistance in water environments". In: *Current opinion in biotechnology* 19.3, pp. 260–265. ISSN: 0958-1669.

Bastyns, K, M Froeyer, Guido Volckaert, and Yves Engelborghs (1994). "The role of Glu-60 in the specificity of the recombinant ribonuclease from *Bacillus amyloliquefaciens* (barnase) towards dinucleotides, poly (A) and RNA." In: *Biochem. J* 300, pp. 737–742.

Beaber, John W., Bianca Hochhut, and Matthew K. Waldor (2004). "SOS response promotes horizontal dissemination of antibiotic resistance genes". In: *Nature* 427.6969, pp. 72–74. ISSN: 0028-0836.

Beaman, Teofila C., A. D. Hitchins, K. Ochi, N. Vasantha, T. Endo, and E. Freese (1983). "Specificity and control of uptake of purines and other compounds in *Bacillus subtilis*". In: *Journal of Bacteriology* 156.3, pp. 1107–1117. ISSN: 0021-9193.

Benveniste, Raoul and Julian Davies (1973). "Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria". In: *Proceedings of the National Academy of Sciences* 70.8, pp. 2276–2280. ISSN: 0027-8424.

Benz, Juliane and Anton Meinhart (2014). "Antibacterial effector/immunity systems: it's just the tip of the iceberg". In: *Current Opinion in Microbiology* 17, pp. 1–10. URL: `http://www.sciencedirect.com/science/article/pii/S1369527413002178`.

Bergstrom, Carl T., Marc Lipsitch, and Bruce R. Levin (2000). "Natural selection, infectious transfer and the existence conditions for bacterial plasmids". In: *Genetics* 155.4, pp. 1505–1519. ISSN: 0016-6731.

Berleman, James E. and John R. Kirby (2009). "Deciphering the hunting strategy of a bacterial wolfpack". In: *FEMS microbiology reviews* 33.5, pp. 942–957. ISSN: 1574-6976.

Bickle, Thomas A. and D. H. Krüger (1993). "Biology of DNA restriction". In: *Microbiological reviews* 57.2, pp. 434–450. ISSN: 1092-2172.

Bingham, Ryan, Stephen I. N. Ekunwe, Sherry Falk, Larry Snyder, and Colin Kleanthous (2000). "The major head protein of bacteriophage T4 binds specifically to elongation factor Tu". In: *Journal of Biological Chemistry* 275.30, pp. 23219–23226. ISSN: 0021-9258.

Birmingham, Amanda, Emily M. Anderson, Angela Reynolds, Diane Ilsley-Tyree, Devin Leake, Yuriy Fedorov, Scott Baskerville, Elena Maksimova, Kathryn Robinson, and Jon Karpilow (2006). "3' UTR seed matches, but not overall identity, are

associated with RNAi off-targets". In: *Nature methods* 3.3, pp. 199–204. ISSN: 1548-7091.

Bisht, N. C., A. Jagannath, P. K. Burma, A. K. Pradhan, and D. Pental (2007). "Retransformation of a male sterile barnase line with the barstar gene as an efficient alternative method to identify male sterile-restorer combinations for heterosis breeding". In: *Plant Cell Reports* 26.6, pp. 727–733.

Blower, Tim R., Terry J. Evans, Rita Przybilski, Peter C. Fineran, and George P. C. Salmond (2012). "Viral evasion of a bacterial suicide system by RNA–based molecular mimicry enables infectious altruism". In: *PLoS genetics* 8.10, e1003023. ISSN: 1553-7404.

Blower, Tim R., Xue Y. Pei, Francesca L. Short, Peter C. Fineran, David P. Humphreys, Ben F. Luisi, and George P. C. Salmond (2011). "A processed noncoding RNA regulates an altruistic bacterial antiviral system". In: *Nature structural & molecular biology* 18.2, pp. 185–190. ISSN: 1545-9993.

Bodmer, Walter F. and Susan Grether (1965). "Uptake and incorporation of thymine, thymidine, uracil, uridine, and 5-fluorouracil into the nucleic acids of *Bacillus subtilis*". In: *Journal of Bacteriology* 89.4, pp. 1011–1014. ISSN: 0021-9193.

Bourniquel, Aude A. and Thomas A. Bickle (2002). "Complex restriction enzymes: NTP-driven molecular motors". In: *Biochimie* 84.11, pp. 1047–1059. ISSN: 0300-9084.

Boyd, E. Fidelma, Brigid M. Davis, and Bianca Hochhut (2001). "Bacteriophage-bacteriophage interactions in the evolution of pathogenic bacteria". In: *Trends in microbiology* 9.3, pp. 137–144. ISSN: 0966-842X.

Brantl, Sabine (2012). "Bacterial type I toxin-antitoxin systems". In: *RNA biology* 9.12, pp. 1488–1490. ISSN: 1547-6286.

Brechtel, Elke and Hubert Bahl (1999). "In *Thermoanaerobacterium thermosulfurigenes* EM1 S-layer homology domains do not attach to peptidoglycan". In: *Journal of Bacteriology* 181.16, pp. 5017–5023. ISSN: 0021-9193.

Brown, Jason M. and Karen Joy Shaw (2003). "A novel family of *Escherichia coli* toxin-antitoxin gene pairs". In: *Journal of Bacteriology* 185.22, pp. 6600–6608. ISSN: 0021-9193.

Brzozowska, Iwona and Urszula Zielenkiewicz (2013). "Regulation of toxin–antitoxin systems by proteolysis". In: *Plasmid* 70.1, pp. 33–41. ISSN: 0147-619X.

Buckle, Ashley M. and Alan R. Fersht (1994). "Subsite Binding in an RNase: Structure of a Barnase-Tetranucleotide Complex at 1.76-. ANG. Resolution". In: *Biochemistry* 33.7, pp. 1644–1653. ISSN: 0006-2960.

Bujnicki, J. M. (2003). "Crystallographic and bioinformatic studies on restriction endonucleases: inference of evolutionary relationships in the" midnight zone" of homology". In: *Current Protein and Peptide Science* 4.5, pp. 327–337. ISSN: 1389-2037.

Butala, Matej, Darja Žgur-Bertok, and Steve J. W. Busby (2009). "The bacterial LexA transcriptional repressor". In: *Cellular and Molecular Life Sciences* 66.1, pp. 82–93. ISSN: 1420-682X.

Buts, L., J. Lah, M. H. Dao-Thi, L. Wyns, and R. Loris (2005). "Toxin-antitoxin modules as bacterial metabolic stress managers". In: *Trends in Biochemical Sciences* 30.12, pp. 672–679.

Cascales, Eric, Susan K. Buchanan, Denis Duché, Colin Kleanthous, Roland Lloubes, Kathleen Postle, Margaret Riley, Stephen Slatin, and Daniele Cavard (2007). "Colicin biology". In: *Microbiology and Molecular Biology Reviews* 71.1, pp. 158–229. ISSN: 1092-2172.

Cataudella, Ilaria, Kim Sneppen, Kenn Gerdes, and Namiko Mitarai (2013). "Conditional cooperativity of toxin-antitoxin regulation can mediate bistability between growth and dormancy". In: *PLoS computational biology* 9.8, e1003174. ISSN: 1553-7358.

Cataudella, Ilaria, Ala Trusina, Kim Sneppen, Kenn Gerdes, and Namiko Mitarai (2012). "Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery". In: *Nucleic Acids Research* 40.14, pp. 6424–6434. ISSN: 0305-1048.

Caulcott, Celia A., Andrew Dunn, Heather A. Robertson, Neill S. Cooper, Michael E. Brown, and P. Malcolm Rhodes (1987). "Investigation of the Effect of Growth Environment on the Stability of Low-copy-number Plasmids in *Escherichia coli*". In: *Journal of general microbiology* 133.7, pp. 1881–1889. ISSN: 1350-0872.

Chen, Mario and Vasantha Nagarajan (1993). "The roles of signal peptide and mature protein in RNase (barnase) export from *Bacillus subtilis*". In: *Molecular and General Genetics MGG* 239.3, pp. 409–415. ISSN: 0026-8925.

Cheng, Xiaogang, WenFang Wang, and Ian J. Molineux (2004). "F exclusion of bacteriophage T7 occurs at the cell membrane". In: *Virology* 326.2, pp. 340–352. ISSN: 0042-6822.

Cherny, Izhack, Martin Overgaard, Jonas Borch, Yaron Bram, Kenn Gerdes, and Ehud Gazit (2007). "Structural and thermodynamic characterization of the *Escherichia coli* RelBE toxin-antitoxin system: indication for a functional role of differential stability". In: *Biochemistry* 46.43, pp. 12152–12163. ISSN: 0006-2960.

Cherny, Izhack, Liat Rockah, and Ehud Gazit (2005). "The YoeB Toxin Is a Folded Protein That Forms a Physical Complex with the Unfolded YefM Antitoxin IMPLICATIONS FOR A STRUCTURAL-BASED DIFFERENTIAL STABILITY OF TOXIN-ANTITOXIN SYSTEMS". In: *Journal of Biological Chemistry* 280.34, pp. 30063–30072. ISSN: 0021-9258.

Chinen, A., Y. Naito, N. Handa, and I. Kobayashi (2000). "Evolution of sequence recognition by restriction-modification enzymes: selective pressure for specificity decrease". In: *Molecular Biology and Evolution* 17.11, pp. 1610–1619. ISSN: 0737-4038.

Chopin, M. C., A. Chopin, and E. Bidnenko (2005). "Phage abortive infection in lactococci: variations on a theme". In: *Current opinion in microbiology* 8.4, pp. 473–479. ISSN: 1369-5274.

Chopra, Nikita, Abhinandan Pathak, Rakesh Bhatnagar, and Sonika Bhatnagar (2013). "Linkage, Mobility, and Selfishness in the MazF Family of Bacterial Toxins: A Snapshot of Bacterial Evolution". In: *Genome biology and evolution* 5.12, pp. 2268–2284. ISSN: 1759-6653.

Christensen, S. K. and K. Gerdes (2003). "RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA". In: *Molecular Microbiology* 48.5, pp. 1389–1400.

Christensen, S. K., G. Maenhaut-Michel, N. Mine, S. Gottesman, K. Gerdes, and L. Van Melderen (2004). "Overproduction of the Lon protease triggers inhibition of translation in *Escherichia coli*: involvement of the yefM-yoeB toxin-antitoxin system". In: *Molecular Microbiology* 51.6, pp. 1705–1717.

Christensen, S. K., M. Mikkelsen, K. Pedersen, and K. Gerdes (2001). "RelE, a global inhibitor of translation, is activated during nutritional stress". In: *Proceedings of the National Academy of Sciences* 98.25, p. 14328.

Christensen, S. K., K. Pedersen, F. G. Hansen, and K. Gerdes (2003). "Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA". In: *Journal of Molecular Biology* 332.4, pp. 809–819.

Christensen-Dalsgaard, M. and K. Gerdes (2006). "Two higBA loci in the *Vibrio cholerae* superintegron encode mRNA cleaving enzymes and can stabilize plasmids". In: *Molecular Microbiology* 62.2, pp. 397–411.

Clark, Alvin J and Gareth J Warren (1979). "Conjugal transmission of plasmids". In: *Annual review of genetics* 13.1, pp. 99–125.

Claverys, J. P., Bernard Martin, and Patrice Polard (2009). "The genetic transformation machinery: composition, localization, and mechanism". In: *FEMS microbiology reviews* 33.3, pp. 643–656. ISSN: 1574-6976.

Claverys, J. P., M. Prudhomme, and B. Martin (2006). "Induction of competence regulons as a general response to stress in Gram-positive bacteria". In: *Annu. Rev. Microbiol.* 60, pp. 451–475. ISSN: 0066-4227.

Condit, R. C. (1975). "F factor-mediated inhibition of bacteriophage T7 growth: Increased membrane permeability and decreased ATP levels following T7 infection of male *Escherichia coli*". In: *Journal of Molecular Biology* 98.1, pp. 45–56. ISSN: 0022-2836.

Condon, C. and H. Putzer (2002). "The phylogenetic distribution of bacterial ribonucleases". In: *Nucleic Acids Research* 30.24, p. 5339.

Cooper, Gregory M and Jay Shendure (2011). "Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data". In: *Nature Reviews Genetics* 12.9, pp. 628–640.

Cooper, T. F. and J. A. Heinemann (2000). "Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids". In: *Proceedings of the National Academy of Sciences of the United States of America* 97.23, p. 12643.

– (2005). "Selection for plasmid post-segregational killing depends on multiple infection: evidence for the selection of more virulent parasites through parasite-level competition". In: *Proceedings of the Royal Society B: Biological Sciences* 272.1561, p. 403. ISSN: 0962-8452.

Correia, F. F., A. D'Onofrio, T. Rejtar, L. Li, B. L. Karger, K. Makarova, E. V. Koonin, and K. Lewis (2006). "Kinase activity of overexpressed HipA is required

for growth arrest and multidrug tolerance in *Escherichia coli*". In: *Journal of Bacteriology* 188.24, p. 8360.

Cram, David, Animesh Ray, and Ron Skurray (1984). "Molecular analysis of F plasmid pif region specifying abortive infection of T7 phage". In: *Molecular and General Genetics MGG* 197.1, pp. 137–142. ISSN: 0026-8925.

Cupples, Claire G. and Jeffrey H. Miller (1989). "A set of lacZ mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions". In: *Proceedings of the National Academy of Sciences* 86.14, pp. 5345–5349. ISSN: 0027-8424.

Dahlberg, Cecilia and Lin Chao (2003). "Amelioration of the cost of conjugative plasmid carriage in Eschericha coli K12". In: *Genetics* 165.4, pp. 1641–1649. ISSN: 0016-6731.

Dao-Thi, Minh-Hoa, Joris Messens, Lode Wyns, and Jan Backmann (2000). "The thermodynamic stability of the proteins of the ccd plasmid addiction system". In: *Journal of Molecular Biology* 299.5, pp. 1373–1386. ISSN: 0022-2836.

Darfeuille, Fabien, Cecilia Unoson, Jørg Vogel, and E. Gerhart H. Wagner (2007). "An antisense RNA inhibits translation by competing with standby ribosomes". In: *Molecular cell* 26.3, pp. 381–392. ISSN: 1097-2765.

Davies, Julian and Dorothy Davies (2010). "Origins and evolution of antibiotic resistance". In: *Microbiology and Molecular Biology Reviews* 74.3, pp. 417–433. ISSN: 1092-2172.

Davison, J. (2005). "Risk mitigation of genetically modified bacteria and plants designed for bioremediation". In: *Journal of Industrial Microbiology and Biotechnology* 32.11, pp. 639–650.

Day, Anthony G., Derek Parsonage, Susanne Ebel, Tom Brown, and Alan R. Fersht (1992). "Barnase has subsites that give rise to large rate enhancements". In: *Biochemistry* 31.28, pp. 6390–6395. ISSN: 0006-2960.

De Bast, Manuel Saavedra, Natacha Mine, and Laurence Van Melderen (2008). "Chromosomal toxin-antitoxin systems may act as antiaddiction modules". In: *Journal of Bacteriology* 190.13, pp. 4603–4609. ISSN: 0021-9193.

Deering, Reginald A, Robert B Guyer, Lenore Stevens, and Theresa E Watson-Thais (1996). "Some repair-deficient mutants of *Dictyostelium discoideum* display enhanced susceptibilities to bleomycin." In: *Antimicrobial agents and chemotherapy* 40.2, pp. 464–467.

Demerec, Milisav and Philip E. Hartman (1959). "Complex loci in microorganisms". In: *Annual Reviews in Microbiology* 13.1, pp. 377–406. ISSN: 0066-4227.

Dempsey, Rita M., David Carroll, Huimin Kong, Lauren Higgins, Conor T. Keane, and David C. Coleman (2005). "Sau42I, a BcgI-like restriction-modification system encoded by the Staphylococcus aureus quadruple-converting phage π42". In: *Microbiology* 151.4, pp. 1301–1311. ISSN: 1350-0872.

Diago-Navarro, Elizabeth, Ana María Hernández-Arriaga, Slawomir Kubik, Igor Konieczny, and Ramón Díaz-Orejas (2013). "Cleavage of the antitoxin of the *parD* toxin-antitoxin system is determined by the ClpAP protease and is modulated by the relative ratio of the toxin and the antitoxin". In: *Plasmid* 70.1, pp. 78–85. ISSN: 0147-619X.

Dörr, Tobias, Kim Lewis, and Marin Vulić (2009). "SOS response induces persistence to fluoroquinolones in *Escherichia coli*". In: *PLoS Genetics* 5.12, e1000760. ISSN: 1553-7404.

Dörr, Tobias, Marin Vulić, and Kim Lewis (2010). "Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*". In: *PLoS biology* 8.2, e1000317. ISSN: 1545-7885.

Duport, Catherine, Christine Baysse, and Yvon Michel-Briand (1995). "Molecular characterization of pyocin S3, a novel S-type pyocin from *Pseudomonas aeruginosa*". In: *Journal of Biological Chemistry* 270.15, pp. 8920–8927. ISSN: 0021-9258.

Dy, Ron L., Rita Przybilski, Koen Semeijn, George P. C. Salmond, and Peter C. Fineran (2014). "A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism". In: *Nucleic Acids Research*, gkt1419. ISSN: 0305-1048.

Eberhard, William G. (1990). "Evolution in bacterial plasmids and levels of selection". In: *Quarterly Review of Biology*, pp. 3–22. ISSN: 0033-5770.

Ebersbach, Gitte and Kenn Gerdes (2005). "Plasmid segregation mechanisms". In: *Annu. Rev. Genet.* 39, pp. 453–479. ISSN: 0066-4197.

Ellstrand, N. C. (2001). "When transgenes wander, should we worry?" In: *Plant physiology* 125.4, p. 1543.

Ellstrand, N. C., H. C. Prentice, and J. F. Hancock (1999). "Gene flow and introgression from domesticated plants into their wild relatives". In: *Annual Review of Ecology and Systematics* 30.1, pp. 539–563.

Elsenstein, F. "PHYLIP-phylogeny inference package (version 3.2)". In: *Cladistics* 5, pp. 164–166.

Engelberg-Kulka, Hanna, Shahar Amitai, Ilana Kolodkin-Gal, and Ronen Hazan (2006). "Bacterial programmed cell death and multicellular behavior in bacteria". In: *PLoS Genetics* 2.10, e135. ISSN: 1553-7404.

Engelberg–Kulka, Hanna, Myriam Reches, Sudarsan Narasimhan, Rachel Schoulaker-Schwarz, Yoel Klemes, Einat Aizenman, and Gad Glaser (1998). "*rex B* of bacteriophage λ is an anti-cell death gene". In: *Proceedings of the National Academy of Sciences* 95.26, pp. 15481–15486. ISSN: 0027-8424.

Enis, Peter and Seymour Geisser (1971). "Estimation of the probability that Y< X". In: *Journal of the American Statistical Association* 66.333, pp. 162–168.

Fasani, Rick A. and Michael A. Savageau (2013). "Molecular mechanisms of multiple toxin-antitoxin systems are coordinated to govern the persister phenotype". In: *Proceedings of the National Academy of Sciences*. ISSN: 0027-8424.

Felsenstein, Joseph (1981). "Evolutionary trees from DNA sequences: a maximum likelihood approach". In: *Journal of molecular evolution* 17.6, pp. 368–376. ISSN: 0022-2844.

Feyter, Rob de, Carolyn Wallace, and David Lane (1989). "Autoregulation of the *ccd* operon in the F plasmid". In: *Molecular and General Genetics MGG* 218.3, pp. 481–486. ISSN: 0026-8925.

Fiebig, Aretha, Cyd Marie Castro Rojas, Dan Siegal-Gaskins, and Sean Crosson (2010). "Interaction specificity, toxicity and regulation of a paralogous set of ParE/RelE-family toxin-antitoxin systems". In: *Molecular Microbiology* 77.1, pp. 236–251. ISSN: 1365-2958.

Findeiß, Sven, Cornelius Schmidtke, Peter F. Stadler, and Ulla Bonas (2010). "A novel family of plasmid-transferred anti-sense ncRNAs". In: *RNA biology* 7.2, pp. 120–124. ISSN: 1547-6286.

Fineran, Peter C., Tim R. Blower, Ian J. Foulds, David P. Humphreys, Kathryn S. Lilley, and George P. C. Salmond (2009). "The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair". In: *Proceedings of the National Academy of Sciences* 106.3, pp. 894–899. ISSN: 0027-8424.

Forsberg, Kevin J., Alejandro Reyes, Bin Wang, Elizabeth M. Selleck, Morten O. A. Sommer, and Gautam Dantas (2012). "The shared antibiotic resistome of soil

bacteria and human pathogens". In: *Science* 337.6098, pp. 1107–1111. ISSN: 0036-8075.

Fozo, Elizabeth M., Matthew R. Hemm, and Gisela Storz (2008). "Small toxic proteins and the antisense RNAs that repress them". In: *Microbiology and Molecular Biology Reviews* 72.4, pp. 579–589. ISSN: 1092-2172.

Fozo, Elizabeth M., Mitsuoki Kawano, Fanette Fontaine, Yusuf Kaya, Kathy S. Mendieta, Kristi L. Jones, Alejandro Ocampo, Kenneth E. Rudd, and Gisela Storz (2008). "Repression of small toxic protein synthesis by the Sib and OhsC small RNAs". In: *Molecular Microbiology* 70.5, pp. 1076–1093. ISSN: 1365-2958.

Fozo, Elizabeth M., Kira S. Makarova, Svetlana A. Shabalina, Natalya Yutin, Eugene V. Koonin, and Gisela Storz (2010). "Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families". In: *Nucleic Acids Research* 38.11, pp. 3743–3759. ISSN: 0305-1048.

Franch, Thomas, Alexander P. Gultyaev, and Kenn Gerdes (1997). "Programmed cell death by *hok/sok* of plasmid R1: Processing at the *hok* mRNA-end triggers structural rearrangements that allow translation and antisense RNA binding". In: *Journal of Molecular Biology* 273.1, pp. 38–51. ISSN: 0022-2836.

Frost, Laura S., Raphael Leplae, Anne O. Summers, and Ariane Toussaint (2005). "Mobile genetic elements: the agents of open source evolution". In: *Nature Reviews Microbiology* 3.9, pp. 722–732. ISSN: 1740-1526.

Furuta, Y., K. Abe, and I. Kobayashi (2010). "Genome comparison and context analysis reveals putative mobile forms of restriction-modification systems and related rearrangements". In: *Nucleic Acids Research* 38.7, pp. 2428–2443. ISSN: 0305-1048.

García-Sogo, Begoña, Benito Pineda, Lourdes Castelblanque, Teresa Antón, Mónica Medina, Edelín Roque, Claudia Torresi, José Pío Beltrán, Vicente Moreno, and Luis Antonio Cañas (2010). "Efficient transformation of *Kalanchoe blossfeldiana* and production of male-sterile plants by engineered anther ablation". In: *Plant cell reports* 29.1, pp. 61–77.

Gazit, Ehud and Robert T. Sauer (1999). "The Doc toxin and Phd antidote proteins of the bacteriophage P1 plasmid addiction system form a heterotrimeric complex". In: *Journal of Biological Chemistry* 274.24, pp. 16813–16818. ISSN: 0021-9258.

Gerdes, K., F. W. Bech, S. T. Jørgensen, A. Løbner-Olesen, P. B. Rasmussen, T. Atlung, L. Boe, O. Karlstrom, S. Molin, and K. Von Meyenburg (1986). "Mechanism

of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the *E. coli* relB operon". In: *The EMBO journal* 5.8, p. 2023.

Gerdes, K., S. K. Christensen, and A. Løbner-Olesen (2005). "Prokaryotic toxin-antitoxin stress response loci". In: *Nature Reviews Microbiology* 3.5, pp. 371–382.

Gerdes, K., T. Thisted, and J. Martinussen (1990). "Mechanism of post-segregational killing by the hoklsok system of plasmid R1: sok antisense RNA regulates formation of a hok mRNA species correlated with killing of plasmid-free cells". In: *Molecular Microbiology* 4.11, pp. 1807–1818. ISSN: 1365-2958.

Gerdes, Kenn (2000). "Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress". In: *Journal of Bacteriology* 182.3, pp. 561–572.

Gerdes, Kenn, Allan Nielsen, Peter Thorsted, and E. Gerhart H. Wagner (1992). "Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable hok, srnB and pndA effector messenger RNAs". In: *Journal of Molecular Biology* 226.3, pp. 637–649. ISSN: 0022-2836.

Gerdes, Kenn, Poul B. Rasmussen, and Søren Molin (1986). "Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells". In: *Proceedings of the National Academy of Sciences* 83.10, pp. 3116–3120. ISSN: 0027-8424.

Gerdes, Kenn and E. Gerhart H. Wagner (2007). "RNA antitoxins". In: *Current opinion in microbiology* 10.2, pp. 117–124. ISSN: 1369-5274.

Goeders, Nathalie and Laurence Van Melderen (2014). "Toxin-Antitoxin Systems as Multilevel Interaction Systems". In: *Toxins* 6.1, pp. 304–324.

González-Candelas, Fernando and M. Pilar Francino (2012). "Barriers to Horizontal Gene Transfer: Fuzzy and Evolvable Boundaries". In: *Horizontal Gene Transfer in Microorganisms*, p. 47.

González-Pastor, José Eduardo (2011). "Cannibalism: a social behavior in sporulating *Bacillus subtilis*". In: *FEMS microbiology reviews* 35.3, pp. 415–424.

Gotfredsen, Marie and Kenn Gerdes (1998). "The *Escherichia coli* relBE genes belong to a new toxin-antitoxin gene family". In: *Molecular Microbiology* 29.4, pp. 1065–1076. ISSN: 1365-2958.

Grady, Ruth and Finbarr Hayes (2003). "Axe-Txe, a broad-spectrum proteic toxin-antitoxin system specified by a multidrug-resistant, clinical isolate of Enterococcus faecium". In: *Molecular Microbiology* 47.5, pp. 1419–1432. ISSN: 1365-2958.

Grant, Seth G., Joel Jessee, Fredric R. Bloom, and Douglas Hanahan (1990). "Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants". In: *Proceedings of the National Academy of Sciences* 87.12, pp. 4645–4649. ISSN: 0027-8424.

Gray, Tamara E., Jorg Eder, Mark Bycroft, Anthony G. Day, and Alan R. Fersht (1993). "Refolding of barnase mutants and pro-barnase in the presence and absence of GroEL". In: *The EMBO journal* 12.11, p. 4145.

Grønlund, Hugo and Kenn Gerdes (1999). "Toxin-antitoxin systems homologous with *relBE* of *Escherichia coli* plasmid P307 are ubiquitous in prokaryotes". In: *Journal of Molecular Biology* 285.4, pp. 1401–1415.

Guasch, J. F., J. Enfedaque, S. Ferrer, D. Gargallo, and M. Regue (1995). "Bacteriocin 28b, a chromosomally encoded bacteriocin produced by most Serratia marcescens biotypes". In: *Research in microbiology* 146.6, pp. 477–483. ISSN: 0923-2508.

Guerout, Anne-Marie, Naeem Iqbal, Natacha Mine, Magaly Ducos-Galand, Laurence Van Melderen, and Didier Mazel (2013). "Characterization of the phd-doc and ccd Toxin-Antitoxin Cassettes from Vibrio Superintegrons". In: *Journal of Bacteriology* 195.10, pp. 2270–2283. ISSN: 0021-9193.

Guglielmini, Julien and Laurence Van Melderen (2012). "Do not distribute". In:

Identification of a sex-factor-affinity site in E. coli as $\gamma\delta$ (1981). Vol. 45. Cold Spring Harbor symposia on quantitative biology. Cold Spring Harbor Laboratory Press, pp. 135–140.

Guzman, Luz-Maria, Dominique Belin, Michael J. Carson, and J. O. N. Beckwith (1995). "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter". In: *Journal of Bacteriology* 177.14, pp. 4121–4130. ISSN: 0021-9193.

Handa, Naofumi, Asao Ichige, Kohji Kusano, and Ichizo Kobayashi (2000). "Cellular responses to postsegregational killing by restriction-modification genes". In: *Journal of Bacteriology* 182.8, pp. 2218–2229. ISSN: 0021-9193.

Hansen, Sonja, Kim Lewis, and Marin Vulić (2008). "Role of global regulators and nucleotide metabolism in antibiotic tolerance in *Escherichia coli*". In: *Antimicrobial agents and chemotherapy* 52.8, pp. 2718–2726. ISSN: 0066-4804.

Hardy, K. G., G. G. Meynell, J. E. Dowman, and B. G. Spratt (1973). "Two major groups of colicin factors: their evolutionary significance". In: *Molecular and General Genetics MGG* 125.3, pp. 217–230. ISSN: 0026-8925.

Harrison, Freya (2013). "Bacterial cooperation in the wild and in the clinic: Are pathogen social behaviours relevant outside the laboratory?" In: *Bioessays* 35.2, pp. 108–112. ISSN: 1521-1878.

Hartley, R. W. and J. R. Smeaton (1973). "On the reaction between the extracellular ribonuclease of *Bacillus amyloliquefaciens* (barnase) and its intracellular inhibitor (barstar)". In: *Journal of Biological Chemistry* 248.16, p. 5624.

Hartley, Robert W (1988). "Barnase and barstar: expression of its cloned inhibitor permits expression of a cloned ribonuclease". In: *Journal of molecular biology* 202.4, pp. 913–915.

Hartley, Robert W. (1993). "Directed mutagenesis and barnase-barstar recognition". In: *Biochemistry* 32.23, pp. 5978–5984. ISSN: 0006-2960.

Hartley, Robert W., Valentin Both, Eric J. Hebert, Dagmar Homerova, Milan Jucovic, Viktor Nazarov, Igor Rybajlak, and Jozef Sevcik (1996). "Barstar inhibits extracellular ribonucleases of Streptomyces and allows their production from recombinant genes". In: *Protein and Peptide Letters* 3.4, pp. 225–231. ISSN: 0929-8665.

Harwood, Colin R and Rocky Cranenburgh (2008). "*Bacillus* protein secretion: an unfolding story". In: *Trends in microbiology* 16.2, pp. 73–79.

Hashimoto-Gotoh, T., F. C. H. Franklin, A. Nordheim, and K. N. Timmis (1981). "Specific-purpose plasmid cloning vectors I. Low copy number, temperature-sensitive, mobilization-defective pSC101-derived containment vectors". In: *Gene* 16.1, pp. 227–235. ISSN: 0378-1119.

Hayes, F. (2003). "Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest". In: *Science* 301.5639, p. 1496.

Hazan, R. and H. Engelberg-Kulka (2004). "*Escherichia coli mazEF*-mediated cell death as a defense mechanism that inhibits the spread of phage P1". In: *Molecular Genetics and Genomics* 272.2, pp. 227–234. ISSN: 1617-4615.

Heinemann, J. A. (1999). "How antibiotics cause antibiotic resistance". In: *Drug discovery today* 4.2, pp. 72–79. ISSN: 1359-6446.

Heinemann, J. A. and M. W. Silby (2003). "Horizontal gene transfer and the selection of antibiotic resistance". In: *Multiple drug resistant bacteria. Horizon, Wymondham*, pp. 161–178.

Heinemann, J. A. and T. Traavik (2004a). "Problems in monitoring horizontal gene transfer in field trials of transgenic plants". In: *Nature Biotechnology* 22.9, pp. 1105–1109.

Heinemann, JA (2007). *A typology of the effects of (trans) gene flow on the conservation and sustainable use of genetic resources. Bsp35rev1. UN FAO.*

Heinemann, Jack A (1991). "Genetics of gene transfer between species". In: *Trends in Genetics* 7.6, pp. 181–185.

Heinemann, Jack A. (1998). "Looking sideways at the evolution of replicons". In: *horizontal gene transfer*, pp. 11–24.

Heinemann, Jack A. and Robert G. Ankenbauer (1993). "Retrotransfer of IncP piasmid R751 from *Escherichia coli* maxicells: evidence for the genetic sufficiency of self–transferable plasmids for bacterial conjugation". In: *Molecular Microbiology* 10.1, pp. 57–62. ISSN: 1365-2958.

Heinemann, Jack A, Robert G Ankenbauer, and Carlos F Amábile-Cuevas (2000). "Do antibiotics maintain antibiotic resistance?" In: *Drug Discovery Today* 5.5, pp. 195–204.

Heinemann, Jack A., Heidi E. Scott, and Meredith Williams (1996). "Doing the conjugative two-step: evidence of recipient autonomy in retrotransfer". In: *Genetics* 143.3, pp. 1425–1435. ISSN: 0016-6731.

Heinemann, Jack A. and George F. Sprague Jr (1989). "Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast". In: *Nature* 340.6230, pp. 205–209. ISSN: 0028-0836.

Heinemann, Jack A. and Terje Traavik (2004b). "Correspondence: Monitoring horizontal gene transfer". In: *Nature Biotechnology* 21, pp. 1349–1350.

Inglis, R. Fredrik, Bihter Bayramoglu, Osnat Gillor, and Martin Ackermann (2013). "The role of bacteriocins as selfish genetic elements". In: *Biology letters* 9.3, p. 20121173. ISSN: 1744-9561.

Isaeva, A. S., E. E. Kulikov, N. V. Ravin, B. D. Dorokhov, K. K. Tarasyan, and A. V. Letarov (2010). "Application of the new *ccdAB*-type natural toxin-antitoxin module for stabilization of inheritance of expressive plasmid vectors based on the bacteriophage N15 replicon in *Escherichia coli* cells". In: *Microbiology* 79.5, pp. 638–645. ISSN: 0026-2617.

Ishikawa, K., E. Fukuda, and I. Kobayashi (2010). "Conflicts targeting epigenetic systems and their resolution by cell death: novel concepts for methyl-specific and other restriction systems". In: *DNA research* 17.6, pp. 325–342. ISSN: 1340-2838.

Ives, C. L., P. D. Nathan, and J. E. Brooks (1992). "Regulation of the BamHI restriction-modification system by a small intergenic open reading frame, bamHIC,

in both *Escherichia coli* and *Bacillus subtilis*". In: *Journal of Bacteriology* 174.22, pp. 7194–7201. ISSN: 0021-9193.

Jackson, Aimee L., Steven R. Bartz, Janell Schelter, Sumire V. Kobayashi, Julja Burchard, Mao Mao, Bin Li, Guy Cavet, and Peter S. Linsley (2003). "Expression profiling reveals off-target gene regulation by RNAi". In: *Nature Biotechnology* 21.6, pp. 635–637. ISSN: 1087-0156.

Jaffe, Aline, T. Ogura, and S. Hiraga (1985). "Effects of the ccd function of the F plasmid on bacterial growth". In: *Journal of Bacteriology* 163.3, pp. 841–849. ISSN: 0021-9193.

Jain, Ravi, Maria C. Rivera, and James A. Lake (1999). "Horizontal gene transfer among genomes: the complexity hypothesis". In: *Proceedings of the National Academy of Sciences* 96.7, pp. 3801–3806. ISSN: 0027-8424.

Jain, Ravi, Maria C. Rivera, Jonathan E. Moore, and James A. Lake (2002). "Horizontal gene transfer in microbial genome evolution". In: *Theoretical population biology* 61.4, pp. 489–495. ISSN: 0040-5809.

Jeong, Haeyoung, Valérie Barbe, Choong Hoon Lee, David Vallenet, Dong Su Yu, Sang-Haeng Choi, Arnaud Couloux, Seung-Won Lee, Sung Ho Yoon, and Laurence Cattolico (2009). "Genome Sequences of *Escherichia coli* B strains REL606 and BL21 (DE3)". In: *Journal of Molecular Biology* 394.4, pp. 644–652. ISSN: 0022-2836.

Jiang, Yong, Joe Pogliano, Donald R. Helinski, and Igor Konieczny (2002). "ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase". In: *Molecular Microbiology* 44.4, pp. 971–979. ISSN: 1365-2958.

Johnson, Erik P., Arne R. Strom, and Donald R. Helinski (1996). "Plasmid RK2 toxin protein ParE: purification and interaction with the ParD antitoxin protein". In: *Journal of Bacteriology* 178.5, pp. 1420–1429. ISSN: 0021-9193.

Kall, Lukas and Erik L. L. Sonnhammer (2002). "Reliability of transmembrane predictions in whole-genome data". In: *Febs Letters* 532.3, pp. 415–418. ISSN: 0014-5793.

Kawano, Mitsuoki (2012). "Divergently overlapping cis-encoded antisense RNA regulating toxin-antitoxin systems from *E. coli*". In: *RNA biology* 9.12, pp. 1520–1527.

Kawano, Mitsuoki, L. Aravind, and Gisela Storz (2007). "An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin". In: *Molecular Microbiology* 64.3, pp. 738–754. ISSN: 1365-2958.

Kawano, Mitsuoki, Taku Oshima, Hiroaki Kasai, and Hirotada Mori (2002). "Molecular characterization of long direct repeat (LDR) sequences expressing a stable mRNA encoding for a 35-amino-acid cell-killing peptide and a cis-encoded small antisense RNA in *Escherichia coli*". In: *Molecular Microbiology* 45.2, pp. 333–349. ISSN: 1365-2958.

Kawano, Mitsuoki, April A. Reynolds, Juan Miranda-Rios, and Gisela Storz (2005). "Detection of 5'-and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*". In: *Nucleic Acids Research* 33.3, pp. 1040–1050. ISSN: 0305-1048.

Kedzierska, Barbara, Lu-Yun Lian, and Finbarr Hayes (2007). "Toxin-antitoxin regulation: bimodal interaction of YefM-YoeB with paired DNA palindromes exerts transcriptional autorepression". In: *Nucleic Acids Research* 35.1, pp. 325–339. ISSN: 0305-1048.

Keren, Iris, Devang Shah, Amy Spoering, Niilo Kaldalu, and Kim Lewis (2004). "Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*". In: *Journal of Bacteriology* 186.24, pp. 8172–8180. ISSN: 0021-9193.

Kharitonova, MA and VI Vershinina (2009). "Biosynthesis of secreted ribonucleases of *Bacillus intermedius* and *Bacillus circulans* under nitrogen starvation". In: *Microbiology* 78.2, pp. 187–192.

Khlebnikov, Artem, Øystein Risa, Tove Skaug, Trent A. Carrier, and J. D. Keasling (2000). "Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture". In: *Journal of Bacteriology* 182.24, pp. 7029–7034. ISSN: 0021-9193.

Khurana, Ritu and Jayant B. Udgaonkar (1994). "Equilibrium unfolding studies of barstar: evidence for an alternative conformation which resembles a molten globule". In: *Biochemistry* 33.1, pp. 106–115. ISSN: 0006-2960.

Kint, Cyrielle I., Natalie Verstraeten, Maarten Fauvart, and Jan Michiels (2012). "New-found fundamentals of bacterial persistence". In: *Trends in microbiology* 20.12, pp. 577–585. ISSN: 0966-842X.

Kita, Keiko, Hideaki Kawakami, and Hiroaki Tanaka (2003). "Evidence for horizontal transfer of the EcoT38I restriction-modification gene to chromosomal DNA by the

P2 phage and diversity of defective P2 prophages in *Escherichia coli* TH38 strains". In: *Journal of Bacteriology* 185.7, pp. 2296–2305. ISSN: 0021-9193.

Klaenhammer, Todd R. (1988). "Bacteriocins of lactic acid bacteria". In: *Biochimie* 70.3, pp. 337–349. ISSN: 0300-9084.

Kleanthous, Colin (2010). "Swimming against the tide: progress and challenges in our understanding of colicin translocation". In: *Nature Reviews Microbiology* 8.12, pp. 843–848. ISSN: 1740-1526.

Knudsen, S. M. and O. H. Karlstrom (1991). "Development of efficient suicide mechanisms for biological containment of bacteria". In: *Applied and Environmental Microbiology* 57.1, p. 85.

Kobayashi, I. (2001). "Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution". In: *Nucleic Acids Research* 29.18, pp. 3742–3756. ISSN: 0305-1048.

Kobayashi, Ichizo (1998). "Selfishness and death: raison d'etre of restriction, recombination and mitochondria". In: *Trends in Genetics* 14.9, pp. 368–374. ISSN: 0168-9525.

Koga, Mitsunori, Yuichi Otsuka, Sébastien Lemire, and Tetsuro Yonesaki (2011). "*Escherichia coli rnlA* and *rnlB* compose a novel toxin–antitoxin system". In: *Genetics* 187.1, pp. 123–130.

Korch, S. B. and T. M. Hill (2006). "Ectopic overexpression of wild-type and mutant *hipA* genes in *Escherichia coli*: effects on macromolecular synthesis and persister formation". In: *Journal of Bacteriology* 188.11, p. 3826.

Korch, Shaleen B., Thomas A. Henderson, and Thomas M. Hill (2003). "Characterization of the *hipA7* allele of *Escherichia coli* and evidence that high persistence is governed by (p) ppGpp synthesis". In: *Molecular Microbiology* 50.4, pp. 1199–1213. ISSN: 1365-2958.

Krajcikova, Daniela and Robert W Hartley (2004). "A new member of the bacterial ribonuclease inhibitor family from *Saccharopolyspora erythraea*". In: *FEBS letters* 557.1, pp. 164–168. URL: `http://www.sciencedirect.com/science/article/pii/S0014579303014686`.

Krajcikova, Daniela, Robert W. Hartley, and Jozef Sevcik (1998). "Isolation and purification of two novel streptomycete RNase inhibitors, SaI14 and SaI20, and cloning, sequencing, and expression in *Escherichia coli* of the gene coding for SaI14". In: *Journal of Bacteriology* 180.6, pp. 1582–1585. ISSN: 0021-9193.

Krogh, Anders, BjoÈrn Larsson, Gunnar Von Heijne, and Erik LL Sonnhammer (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes". In: *Journal of molecular biology* 305.3, pp. 567–580.

Kusano, K., T. Naito, N. Handa, and I. Kobayashi (1995). "Restriction-modification systems as genomic parasites in competition for specific sequences". In: *Proceedings of the National Academy of Sciences* 92.24, p. 11095. ISSN: 0027-8424.

Kwong, Stephen M., Slade O. Jensen, and Neville Firth (2010). "Prevalence of Fst-like toxin-antitoxin systems". In: *Microbiology* 156.4, pp. 975–977. ISSN: 1350-0872.

Labrie, S. J. and S. Moineau (2007). "Abortive infection mechanisms and prophage sequences significantly influence the genetic makeup of emerging lytic lactococcal phages". In: *Journal of Bacteriology* 189.4, p. 1482. ISSN: 0021-9193.

Labrie, Simon J., Julie E. Samson, and Sylvain Moineau (2010). "Bacteriophage resistance mechanisms". In: *Nature Reviews Microbiology* 8.5, pp. 317–327. ISSN: 1740-1526.

Laskaris, Paris, Sahar Tolba, Leo Calvo-Bado, and Liz Wellington (2010). "Coevolution of antibiotic production and counter-resistance in soil bacteria". In: *Environmental microbiology* 12.3, pp. 783–796. ISSN: 1462-2920.

Lawrence, Jeffrey G. and John R. Roth (1996). "Selfish operons: horizontal transfer may drive the evolution of gene clusters". In: *Genetics* 143.4, pp. 1843–1860. ISSN: 0016-6731.

Lee, Sung Kuk, Howard H. Chou, Brian F. Pfleger, Jack D. Newman, Yasuo Yoshikuni, and Jay D. Keasling (2007). "Directed evolution of AraC for improved compatibility of arabinose-and lactose-inducible promoters". In: *Applied and Environmental Microbiology* 73.18, pp. 5711–5715. ISSN: 0099-2240.

Lehnherr, Hansjörg, Emmanuelle Maguin, Samina Jafri, and Michael B Yarmolinsky (1993). "Plasmid Addiction Genes of Bacteriophage P1: *doc*, which Causes Cell Death on Curing of Prophage, and *phd*, which Prevents Host Death when Prophage is Retained". In: *Journal of molecular biology* 233.3, pp. 414–428.

Leipe, Detlef D, L Aravind, and Eugene V Koonin (1999). "Did DNA replication evolve twice independently?" In: *Nucleic acids research* 27.17, pp. 3389–3401. URL: `http://nar.oxfordjournals.org/content/27/17/3389.short;%20http://nar.oxfordjournals.org/content/27/17/3389.full`.

Lemos, José A. C., Thomas A. Brown, Jacqueline Abranches, and Robert A. Burne (2005). "Characteristics of Streptococcus mutans strains lacking the MazEF and RelBE toxin-antitoxin modules". In: *FEMS microbiology letters* 253.2, pp. 251–257. ISSN: 1574-6968.

Lenski, Richard E., Judith A. Mongold, Paul D. Sniegowski, Michael Travisano, Farida Vasi, Philip J. Gerrish, and Thomas M. Schmidt (1998). "Evolution of competitive fitness in experimental populations of *E. coli*: what makes one genotype a better competitor than another?" In: *Antonie van Leeuwenhoek* 73.1, pp. 35–47. ISSN: 0003-6072.

Leplae, Raphaël, Damien Geeraerts, Régis Hallez, Julien Guglielmini, Pierre Drèze, and Laurence Van Melderen (2011). "Diversity of bacterial type II toxin–antitoxin systems: a comprehensive search and functional analysis of novel families". In: *Nucleic acids research* 39.13, pp. 5513–5525.

Lewis, K. (2005). "Persister cells and the riddle of biofilm survival". In: *Biochemistry (Moscow)* 70.2, pp. 267–274.

Li, Ping, Jon Beckwith, and Hiroshi Inouye (1988). "Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*". In: *Proceedings of the National Academy of Sciences* 85.20, pp. 7685–7689. ISSN: 0027-8424.

Liu, Mohan, Yonglong Zhang, Masayori Inouye, and Nancy A. Woychik (2008). "Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit". In: *Proceedings of the National Academy of Sciences* 105.15, pp. 5885–5890. ISSN: 0027-8424.

Lobell, Robert B. and Robert F. Schleif (1990). "DNA looping and unlooping by AraC protein". In: *Science* 250.4980, pp. 528–532. ISSN: 0036-8075.

López, Elena, Marina Elez, Ivan Matic, and Jesús Blázquez (2007). "Antibiotic–mediated recombination: ciprofloxacin stimulates SOS–independent recombination of divergent sequences in *Escherichia coli*". In: *Molecular Microbiology* 64.1, pp. 83–93. ISSN: 1365-2958.

Lorenz, Michael G and Wilfried Wackernagel (1994). "Bacterial gene transfer by natural genetic transformation in the environment." In: *Microbiological reviews* 58.3, p. 563.

Lubienski, Michael J., Mark Bycroft, Stefan M. V. Freund, and Alan R. Fersht (1994). "Three-dimensional solution structure and 13C assignments of barstar using nuclear

magnetic resonance spectroscopy". In: *Biochemistry* 33.30, pp. 8866–8877. ISSN: 0006-2960.

Madera, Martin (2008). "Profile Comparer: a program for scoring and aligning profile hidden Markov models". In: *Bioinformatics* 24.22, pp. 2630–2631. ISSN: 1367-4803.

Magnuson, Roy and Michael B. Yarmolinsky (1998). "Corepression of the P1 addiction operon by Phd and Doc". In: *Journal of Bacteriology* 180.23, pp. 6342–6351. ISSN: 0021-9193.

Maiques, Elisa, Carles Úbeda, Susana Campoy, Noelia Salvador, Íñigo Lasa, Richard P Novick, Jordi Barbé, and José R Penadés (2006). "β-Lactam antibiotics induce the SOS response and horizontal transfer of virulence factors in Staphylococcus aureus". In: *Journal of bacteriology* 188.7, pp. 2726–2729.

Makarova, Kira S., Yuri I. Wolf, and Eugene V. Koonin (2009). "Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes". In: *Biology Direct* 4.1, p. 19. ISSN: 1745-6150.

Maki, Satoko, Souichi Takiguchi, Tadao Horiuchi, Kazuhisa Sekimizu, and Takeyoshi Miki (1996). "Partner Switching Mechanisms in Inactivation and Rejuvenation of *Escherichia coli* DNA Gyrase by F Plasmid Proteins LetD (CcdB) and LetA (CcdA)". In: *Journal of Molecular Biology* 256.3, pp. 473–482. ISSN: 0022-2836.

Maloney, Peter C. and Boris Rotman (1973). "Distribution of suboptimally induced β-d-galactosidase in *Escherichia coli*: The enzyme content of individual cells". In: *Journal of Molecular Biology* 73.1, pp. 77–91. ISSN: 0022-2836.

Mariani, Celestina, Marc De Beuckeleer, Jessie Truettner, Jan Leemans, and Robert B. Goldberg (1990). "Induction of male sterility in plants by a chimaeric ribonuclease gene". In: *Nature* 347.6295, pp. 737–741. ISSN: 0028-0836.

Marianovsky, Irina, Einat Aizenman, Hanna Engelberg-Kulka, and Gad Glaser (2001). "The regulation of the *Escherichia coli mazEF* promoter involves an unusual alternating palindrome". In: *Journal of Biological Chemistry* 276.8, pp. 5975–5984.

Martinez, Jose L. (2012). "Natural antibiotic resistance and contamination by antibiotic resistance determinants: the two ages in the evolution of resistance to antimicrobials". In: *Frontiers in microbiology* 3.

Masuda, Hisako, Qian Tan, Naoki Awano, Kuen-Phon Wu, and Masayori Inouye (2012). "YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ,

antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*". In: *Molecular Microbiology* 84.5, pp. 979–989. ISSN: 1365-2958.

Mathews, David H., Jeffrey Sabina, Michael Zuker, and Douglas H. Turner (1999). "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure". In: *Journal of Molecular Biology* 288.5, pp. 911–940. ISSN: 0022-2836.

Matuschek, Markus, Kerstin Sahm, Arne Zibat, and Hubert Bahl (1996). "Characterization of genes from *Thermoanaerobacterium thermosulfurigenes* EM1 that encode two glycosyl hydrolases with conserved S-layer-like domains". In: *Molecular and General Genetics MGG* 252.4, pp. 493–496. ISSN: 0026-8925.

Mc Ginty, Sorcha E. and Daniel J. Rankin (2012). "The evolution of conflict resolution between plasmids and their bacterial hosts". In: *Evolution* 66.5, pp. 1662–1670. ISSN: 1558-5646.

Melderen, Laurence, Philippe Bernard, and Martine Couturier (1994). "Lon–dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid–free segregant bacteria". In: *Molecular Microbiology* 11.6, pp. 1151–1157. ISSN: 1365-2958.

Michaelis, S., H. Inouye, D. Oliver, and J. Beckwith (1983). "Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*". In: *Journal of bacteriology* 154.1, pp. 366–374. ISSN: 0021-9193.

Miller, Jeffrey H. (1992). *A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria*. Vol. 1. Cold spring harbor laboratory press New York.

Milunovic, Branislava, Richard A. Morton, and Turlough M. Finan (2014). "Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of Sinorhizobium meliloti". In: *Journal of Bacteriology* 196.4, pp. 811–824. ISSN: 0021-9193.

Mine, Natacha, Julien Guglielmini, Myriam Wilbaux, and Laurence Van Melderen (2009). "The decay of the chromosomally encoded ccdO157 toxin-antitoxin system in the *Escherichia coli* species". In: *Genetics* 181.4, pp. 1557–1566. ISSN: 0016-6731.

Mok, Wendy W. K., Nirav H. Patel, and Yingfu Li (2010). "Decoding Toxicity DEDUCING THE SEQUENCE REQUIREMENTS OF IbsC, A TYPE I TOXIN IN ESCHERICHIA COLI". In: *Journal of Biological Chemistry* 285.53, pp. 41627–41636. ISSN: 0021-9258.

Mossakowska, D. E., K. Nyberg, and A. R. Fersht (1989). "Kinetic characterization of the recombinant ribonuclease from *Bacillus amyloliquefaciens* (barnase) and investigation of key residues in catalysis by site-directed mutagenesis". In: *Biochemistry* 28.9, pp. 3843–3850.

Moyed, Harris S. and Kevin P. Bertrand (1983). "*hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis". In: *Journal of Bacteriology* 155.2, pp. 768–775. ISSN: 0021-9193.

Mruk, Iwona and Robert M. Blumenthal (2008). "Real-time kinetics of restriction-modification gene expression after entry into a new host cell". In: *Nucleic Acids Research* 36.8, pp. 2581–2593. ISSN: 0305-1048.

Mruk, Iwona and Ichizo Kobayashi (2014). "To be or not to be: regulation of restriction-modification systems and other toxin-antitoxin systems". In: *Nucleic Acids Research* 42.1, pp. 70–86. ISSN: 0305-1048.

Mulec, Janez, Zdravko Podlesek, Peter Mrak, Andreja Kopitar, Alojz Ihan, and Darja Žgur-Bertok (2003). "A *cka-gfp* transcriptional fusion reveals that the colicin K activity gene is induced in only 3 percent of the population". In: *Journal of bacteriology* 185.2, pp. 654–659.

Naderer, Marc, Jessica R. Brust, Dieter Knowle, and Robert M. Blumenthal (2002). "Mobility of a restriction-modification system revealed by its genetic contexts in three hosts". In: *Journal of Bacteriology* 184.9, pp. 2411–2419. ISSN: 0021-9193.

Naito, T., K. Kusano, and I. Kobayashi (1995). "Selfish behavior of restriction-modification systems". In: *Science* 267.5199, p. 897. ISSN: 0036-8075.

Nakayama, Y. and I. Kobayashi (1998). "Restriction-modification gene complexes as selfish gene entities: roles of a regulatory system in their establishment, maintenance, and apoptotic mutual exclusion". In: *Proceedings of the National Academy of Sciences* 95.11, p. 6442. ISSN: 0027-8424.

Nariya, Hirofumi and Masayori Inouye (2008). "MazF, an mRNA Interferase, Mediates Programmed Cell Death during Multicellular Myxococcus Development". In: *Cell* 132.1, pp. 55–66. ISSN: 0092-8674.

Neubauer, Cajetan, Yong-Gui Gao, Kasper R. Andersen, Christine M. Dunham, Ann C. Kelley, Jendrik Hentschel, Kenn Gerdes, V. Ramakrishnan, and Ditlev E. Brodersen (2009). "The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE". In: *Cell* 139.6, pp. 1084–1095. ISSN: 0092-8674.

Nobusato, Ayaka, Ikuo Uchiyama, and Ichizo Kobayashi (2000). "Diversity of restriction-modification gene homologues in Helicobacter pylori". In: *Gene* 259.1, pp. 89–98. ISSN: 0378-1119.

Nogueira, Teresa, Daniel J. Rankin, Marie Touchon, François Taddei, Sam P. Brown, and Eduardo P. C. Rocha (2009). "Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence". In: *Current Biology* 19.20, pp. 1683–1691. ISSN: 0960-9822.

Noller, Harry F. (2004). "The driving force for molecular evolution of translation". In: *RNA* 10.12, pp. 1833–1837. ISSN: 1355-8382.

Normark, B. Henriques and S. Normark (2002). "Evolution and spread of antibiotic resistance". In: *Journal of internal medicine* 252.2, pp. 91–106. ISSN: 1365-2796.

Novick, Aaron and Milton Weiner (1957). "Enzyme induction as an all-or-none phenomenon". In: *Proceedings of the National Academy of Sciences of the United States of America* 43.7, p. 553.

Nowak, Martin A. and Robert M. May (1994). "Superinfection and the evolution of parasite virulence". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 255.1342, pp. 81–89. ISSN: 0962-8452.

Ochman, Howard, Jeffrey G. Lawrence, and Eduardo A. Groisman (2000). "Lateral gene transfer and the nature of bacterial innovation". In: *Nature* 405.6784, pp. 299–304. ISSN: 0028-0836.

Ogura, Teru and Sota Hiraga (1983). "Mini-F plasmid genes that couple host cell division to plasmid proliferation". In: *Proceedings of the National Academy of Sciences* 80.15, pp. 4784–4788. ISSN: 0027-8424.

O'Neill, Mary, Angela Chen, and Noreen E. Murray (1997). "The restriction-modification genes of *Escherichia coli* K-12 may not be selfish: They do not resist loss and are readily replaced by alleles conferring different specificities". In: *Proceedings of the National Academy of Sciences* 94.26, pp. 14596–14601. ISSN: 0027-8424.

Ono, Tsuneko, Shigeru Akimoto, Kenji Ono, and Yoshinari Ohnishi (1986). "Plasmid genes increase membrane permeability in *Escherichia coli*". In: *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 867.3, pp. 81–88. ISSN: 0167-4781.

Orlowski, J. and J. M. Bujnicki (2008). "Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses". In: *Nucleic Acids Research* 36.11, pp. 3552–3569. ISSN: 0305-1048.

O'Sullivan, David, Denis P. Twomey, Aidan Coffey, Colin Hill, Gerald F. Fitzgerald, and R. Paul Ross (2000). "Novel type I restriction specificities through domain shuffling of HsdS subunits in Lactococcus lactis". In: *Molecular Microbiology* 36.4, pp. 866–875. ISSN: 1365-2958.

Otsuka, Yuichi and Tetsuro Yonesaki (2012). "Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins". In: *Molecular Microbiology* 83.4, pp. 669–681. ISSN: 1365-2958.

Overgaard, Martin, Jonas Borch, Mikkel G. Jørgensen, and Kenn Gerdes (2008). "Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity". In: *Molecular Microbiology* 69.4, pp. 841–857. ISSN: 1365-2958.

Pace, C. Nick, Douglas V. Laurents, and Rick E. Erickson (1992). "Urea denaturation of barnase: pH dependence and characterization of the unfolded state". In: *Biochemistry* 31.10, pp. 2728–2734. ISSN: 0006-2960.

Paddon, C. J., N. Vasantha, and R. W. Hartley (1989). "Translation and processing of *Bacillus amyloliquefaciens* extracellular RNase". In: *Journal of Bacteriology* 171.2, p. 1185.

Paddon, CJ and RW Hartley (1987). "Expression of *Bacillus amyloliquefaciens* extracellular ribonuclease (barnase) in *Escherichia coli* following an inactivating mutation". In: *Gene* 53.1, pp. 11–19.

Pandey, D. P. and K. Gerdes (2005). "Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes". In: *Nucleic Acids Research* 33.3, p. 966.

Parma, David H., Marlene Snyder, Sergey Sobolevski, Myra Nawroz, Edward Brody, and Larry Gold (1992). "The Rex system of bacteriophage lambda: tolerance and altruistic cell death". In: *Genes & development* 6.3, pp. 497–510. ISSN: 0890-9369.

Patel, S. and K. E. Weaver (2006). "Addiction toxin Fst has unique effects on chromosome segregation and cell division in *Enterococcus faecalis* and *Bacillus subtilis*". In: *Journal of Bacteriology* 188.15, pp. 5374–5384. ISSN: 0021-9193.

Pecota, Douglas C, George Osapay, Michael E Selsted, and Thomas K Wood (2003). "Antimicrobial properties of the *Escherichia coli* R1 plasmid host killing peptide". In: *Journal of biotechnology* 100.1, pp. 1–12.

Pecota, Douglas C. and Thomas K. Wood (1996). "Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1". In: *Journal of Bacteriology* 178.7, pp. 2044–2050. ISSN: 0021-9193.

Pedersen, Kim and Kenn Gerdes (1999). "Multiple *hok* genes on the chromosome of *Escherichia coli*". In: *Molecular Microbiology* 32.5, pp. 1090–1102. ISSN: 1365-2958.

Pedersen, Kim, Andrey V. Zavialov, Michael Yu Pavlov, Johan Elf, Kenn Gerdes, and Måns Ehrenberg (2003). "The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site". In: *Cell* 112.1, pp. 131–140. ISSN: 0092-8674.

Pennington, Jeanine M. and Susan M. Rosenberg (2007). "Spontaneous DNA breakage in single living *Escherichia coli* cells". In: *Nature genetics* 39.6, pp. 797–802. ISSN: 1061-4036.

Peters, Joseph E. and Spencer A. Benson (1995). "Redundant transfer of F'plasmids occurs between *Escherichia coli* cells during nonlethal selections". In: *Journal of Bacteriology* 177.3, pp. 847–850. ISSN: 0021-9193.

Pimentel, Belén, Mark A Madine, and Guillermo de la Cueva-Méndez (2005). "Kid cleaves specific mRNAs at UUACU sites to rescue the copy number of plasmid R1". In: *The EMBO journal* 24.19, pp. 3459–3469. ISSN: 1460-2075.

Pimentel, Belén, Radhika Nair, Camino Bermejo-Rodríguez, Mark A Preston, Chukwuma A Agu, Xindan Wang, Juan A Bernal, David J Sherratt, and Guillermo de la Cueva-Méndez (2014). "Toxin Kid uncouples DNA replication and cell division to enforce retention of plasmid R1 in *Escherichia coli* cells". In: *Proceedings of the National Academy of Sciences*, p. 201308241. ISSN: 0027-8424.

Pingoud, Alfred and Albert Jeltsch (2001). "Structure and function of type II restriction endonucleases". In: *Nucleic Acids Research* 29.18, pp. 3705–3727. ISSN: 0305-1048.

Poirel, Laurent, Jose-Manuel Rodriguez-Martinez, Hedi Mammeri, Alain Liard, and Patrice Nordmann (2005). "Origin of plasmid-mediated quinolone resistance determinant QnrA". In: *Antimicrobial agents and chemotherapy* 49.8, pp. 3523–3525. ISSN: 0066-4804.

Poole, Anthony M. and Derek T. Logan (2005). "Modern mRNA proofreading and repair: clues that the last universal common ancestor possessed an RNA genome?" In: *Molecular biology and evolution* 22.6, pp. 1444–1455. ISSN: 0737-4038.

Poulsen, L. K., N. W. Larsen, S. Molin, and P. Andersson (1989). "A family of genes encoding a cell-killing function may be conserved in all Gram-negative bacteria". In: *Molecular Microbiology* 3.11, pp. 1463–1472. ISSN: 1365-2958.

Prakash-Cheng, Ainu, Stephen S. Chung, and Jun-ichi Ryu (1993). "The expression and regulation of hsd K genes after conjugative transfer". In: *Molecular and General Genetics MGG* 241.5-6, pp. 491–496. ISSN: 0026-8925.

Prasun, P, M Pradhan, S Agarwal, *et al.* (2007). "One gene, many phenotypes". In: *Journal of postgraduate medicine* 53.4, p. 257.

Prudhomme, Marc, Laetitia Attaiech, Guillaume Sanchez, Bernard Martin, and Jean-Pierre Claverys (2006). "Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*". In: *Science* 313.5783, pp. 89–92. ISSN: 0036-8075.

Pukall, R, H Tschäpe, and K Smalla (1996). "Monitoring the spread of broad host and narrow host range plasmids in soil microcosms". In: *FEMS microbiology ecology* 20.1, pp. 53–66.

Quinn, Charles T and Scott T Miller (2004). "Risk factors and prediction of outcomes in children and adolescents who have sickle cell anemia". In: *Hematology/oncology clinics of North America* 18.6, pp. 1339–1354.

Raghavendra, Nidhanapati K. and Desirazu N. Rao (2004). "Unidirectional translocation from recognition site and a necessary interaction with DNA end for cleavage by Type III restriction enzyme". In: *Nucleic Acids Research* 32.19, pp. 5703–5711. ISSN: 0305-1048.

Ramage, Holly R., Lynn E. Connolly, and Jeffery S. Cox (2009). "Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution". In: *PLoS Genetics* 5.12, e1000767. ISSN: 1553-7404.

Ramos, Humberto JO, Emanuel M Souza, Juliana RL Soares-Ramos, and Fábio O Pedrosa (2005). "A new system to control the *barnase* expression by a NifA-dependent promoter". In: *Journal of biotechnology* 118.1, pp. 9–16.

– (2006). "Antibiosis by *Bacillus amyloliquefaciens* ribonuclease barnase expressed in *Escherichia coli* against symbiotic and endophytic nitrogen-fixing bacteria". In: *Journal of biotechnology* 126.3, pp. 291–294.

Rankin, D. J., E. P. C. Rocha, and S. P. Brown (2010). "What traits are carried on mobile genetic elements, and why". In: *Heredity* 106.1, pp. 1–10. ISSN: 0018-067X.

Rankin, Daniel J., Leighton A. Turner, Jack A. Heinemann, and Sam P. Brown (2012). "The coevolution of toxin and antitoxin genes drives the dynamics of

bacterial addiction complexes and intragenomic conflict". In: *Proceedings of the Royal Society B: Biological Sciences* 279.1743, pp. 3706–3715. ISSN: 0962-8452.

Rao, Christopher V., Denise M. Wolf, and Adam P. Arkin (2002). "Control, exploitation and tolerance of intracellular noise". In: *Nature* 420.6912, pp. 231–237. ISSN: 0028-0836.

Raser, Jonathan M. and Erin K. O'Shea (2005). "Noise in gene expression: origins, consequences, and control". In: *Science* 309.5743, pp. 2010–2013. ISSN: 0036-8075.

Recorbet, Ghislaine, C. Robert, A. Givaudan, B. Kudla, P. Normand, and G. Faurie (1993). "Conditional suicide system of *Escherichia coli* released into soil that uses the *Bacillus subtilis sacB* gene". In: *Applied and environmental microbiology* 59.5, pp. 1361–1366. ISSN: 0099-2240.

Rensing, Christopher, Deborah T. Newby, and Ian L. Pepper (2002). "The role of selective pressure and selfish DNA in horizontal gene transfer and soil microbial community adaptation". In: *Soil Biology and Biochemistry* 34.3, pp. 285–296. ISSN: 0038-0717.

Riley, Margaret A. and John E. Wertz (2002). "Bacteriocin diversity: ecological and evolutionary perspectives". In: *Biochimie* 84.5, pp. 357–364. ISSN: 0300-9084.

Rivera, Maria C., Ravi Jain, Jonathan E. Moore, and James A. Lake (1998). "Genomic evidence for two functionally distinct gene classes". In: *Proceedings of the National Academy of Sciences* 95.11, pp. 6239–6244. ISSN: 0027-8424.

Roberts, R. J., M. Belfort, T. Bestor, A. S. Bhagwat, T. A. Bickle, J. Bitinaite, R. M. Blumenthal, S. K. Degtyarev, D. T. F. Dryden, and K. Dybvig (2003). "A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes". In: *Nucleic Acids Research* 31.7, pp. 1805–1812. ISSN: 0305-1048.

Roberts, Richard J., Tamas Vincze, Janos Posfai, and Dana Macelis (2007). "REBASE-enzymes and genes for DNA restriction and modification". In: *Nucleic Acids Research* 35.suppl 1, pp. D269–D270. ISSN: 0305-1048.

Rowe-Magnus, Dean A., Anne-Marie Guerout, Latefa Biskri, Philippe Bouige, and Didier Mazel (2003). "Comparative analysis of superintegrons: engineering extensive genetic diversity in the Vibrionaceae". In: *Genome research* 13.3, pp. 428–442. ISSN: 1088-9051.

Rowe-Magnus, Dean A. and Didier Mazel (2001). "Integrons: natural tools for bacterial genome evolution". In: *Current opinion in microbiology* 4.5, pp. 565–569. ISSN: 1369-5274.

Rushizky, G. W., A. E. Greco, R. W. Hartley Jr, and H. A. Sober (1963). "Studies on *B. subtilis* Ribonuclease. I. Characterization of Enzymatic Specificity". In: *Biochemistry* 2.4, pp. 787–793. ISSN: 0006-2960.

Saavedra De Bast, M., N. Mine, and L. Van Melderen (2008). "Chromosomal toxin-antitoxin systems may act as antiaddiction modules". In: *Journal of Bacteriology* 190.13, p. 4603.

Samson, Julie E., Maxime Bélanger, and Sylvain Moineau (2013). "Effect of the abortive infection mechanism and type III toxin/antitoxin system AbiQ on the lytic cycle of *Lactococcus lactis phages*". In: *Journal of Bacteriology* 195.17, pp. 3947–3956. ISSN: 0021-9193.

Santos Sierra, Sandra, Rafael Giraldo, and Ramón Díaz Orejas (1998). "Functional interactions between *chpB* and *parD*, two homologous conditional killer systems found in the *Escherichia coli* chromosome and in plasmid R1". In: *FEMS microbiology letters* 168.1, pp. 51–58. ISSN: 1574-6968.

Sára, Margit and Uwe B. Sleytr (2000). "S-layer proteins". In: *Journal of Bacteriology* 182.4, pp. 859–868. ISSN: 0021-9193.

Sat, Boaz, Ronen Hazan, Tova Fisher, Hanita Khaner, Gad Glaser, and Hanna Engelberg-Kulka (2001). "Programmed Cell Death in *Escherichia coli*: Some Antibiotics Can Trigger *mazE* FLethality". In: *Journal of Bacteriology* 183.6, pp. 2041–2045. ISSN: 0021-9193.

Saxild, Henrik Hans and Per Nygaard (1991). "Regulation of levels of purine biosynthetic enzymes in *Bacillus subtilis*: effects of changing purine nucleotide pools". In: *Journal of general microbiology* 137.10, pp. 2387–2394. ISSN: 1350-0872.

Sayed, Nour, Ambre Jousselin, and Brice Felden (2011). "A cis-antisense RNA acts in trans in *Staphylococcus aureus* to control translation of a human cytolytic peptide". In: *Nature structural & molecular biology* 19.1, pp. 105–112. ISSN: 1545-9993.

Sayed, Nour, Sylvie Nonin-Lecomte, Stphane Rèty, and Brice Felden (2012). "Functional and structural insights of a *Staphylococcus aureus* apoptotic-like membrane peptide from a toxin-antitoxin module". In: *Journal of Biological Chemistry* 287.52, pp. 43454–43463. ISSN: 0021-9258.

Sberro, Hila, Azita Leavitt, Ruth Kiro, Eugene Koh, Yoav Peleg, Udi Qimron, and Rotem Sorek (2013). "Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning". In: *Molecular cell* 50.1, pp. 136–148. ISSN: 1097-2765.

Schleif, Robert (2010). "AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action". In: *FEMS microbiology reviews* 34.5, pp. 779–796. ISSN: 1574-6976.

Schreiber, Gideon, Christian Frisch, and Alan R Fersht (1997). "The role of Glu73 of barnase in catalysis and the binding of barstar". In: *Journal of molecular biology* 270.1, pp. 111–122.

Schuch, Raymond, Araik Garibian, Hans H. Saxild, Patrick J. Piggot, and Per Nygaard (1999). "Nucleosides as a carbon source in *Bacillus subtilis*: characterization of the *drm-pupG* operon". In: *Microbiology* 145.10, pp. 2957–2966. ISSN: 1350-0872.

Schumacher, Maria A., Kevin M. Piro, Weijun Xu, Sonja Hansen, Kim Lewis, and Richard G. Brennan (2009). "Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB". In: *Science* 323.5912, pp. 396–401. ISSN: 0036-8075.

Serrano, Luis, Andreas Matouschek, and Alan R. Fersht (1992). "The folding of an enzyme: III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure". In: *Journal of Molecular Biology* 224.3, pp. 805–818. ISSN: 0022-2836.

Shah, Devang, Zhigang Zhang, Arkady Khodursky, Niilo Kaldalu, Kristi Kurg, and Kim Lewis (2006). "Persisters: a distinct physiological state of *E. coli*". In: *Bmc Microbiology* 6.1, p. 53. ISSN: 1471-2180.

Sharipova, MR, LV Lopukhov, OA Vershinina, and IB Leshchinskaya (2005). "Some secretion characteristics of bacterial ribonuclease". In: *Microbiology* 74.1, pp. 27–31. URL: `http://link.springer.com/article/10.1007/s11021-005-0024-9`.

Short, Francesca L., Xue Y. Pei, Tim R. Blower, Shue-Li Ong, Peter C. Fineran, Ben F. Luisi, and George P. C. Salmond (2013). "Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot". In: *Proceedings of the National Academy of Sciences* 110.3, E241–E249. ISSN: 0027-8424.

Siegele, Deborah A. and James C. Hu (1997). "Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents

mixed populations". In: *Proceedings of the National Academy of Sciences* 94.15, pp. 8168–8172. ISSN: 0027-8424.

Silvaggi, Jessica M., John B. Perkins, and Richard Losick (2005). "Small untranslated RNA antitoxin in *Bacillus subtilis*". In: *Journal of Bacteriology* 187.19, pp. 6641–6650. ISSN: 0021-9193.

Sistla, Srivani and Desirazu N. Rao (2004). "S-Adenosyl-L-methionine-dependent restriction enzymes". In: *Critical reviews in biochemistry and molecular biology* 39.1, pp. 1–19. ISSN: 1040-9238.

Slavcev, R. A. and S. Hayes (2003). "Stationary phase-like properties of the bacteriophage λ Rex exclusion phenotype". In: *Molecular Genetics and Genomics* 269.1, pp. 40–48. ISSN: 1617-4615.

Smith, David L., Jonathan Dushoff, Eli N. Perencevich, Anthony D. Harris, and Simon A. Levin (2004). "Persistent colonization and the spread of antibiotic resistance in nosocomial pathogens: resistance is a regional problem". In: *Proceedings of the National Academy of Sciences of the United States of America* 101.10, pp. 3709–3714. ISSN: 0027-8424.

Solovyev, V. and A. Salamov (2011). "Automatic annotation of microbial genomes and metagenomic sequences". In: *Metagenomics and its applications in agriculture, biomedicine and environmental studies*, pp. 61–78.

Starikova, Irina, Mohammed Al-Haroni, Guido Werner, Adam P. Roberts, Vidar Sørum, Kaare M. Nielsen, and Pål J. Johnsen (2013). "Fitness costs of various mobile genetic elements in *Enterococcus faecium* and *Enterococcus faecalis*". In: *Journal of Antimicrobial Chemotherapy* 68.12, pp. 2755–2765. ISSN: 0305-7453.

Stenberg, G. and A. R. Fersht (1997). "Folding of barnase in the presence of the molecular chaperone SecB". In: *Journal of Molecular Biology* 274.2, pp. 268–275.

Studier, F. William and Pradip K. Bandyopadhyay (1988). "Model for how type I restriction enzymes select cleavage sites in DNA". In: *Proceedings of the National Academy of Sciences* 85.13, pp. 4677–4681. ISSN: 0027-8424.

Sundin, George W. (2007). "Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts". In: *Annu. Rev. Phytopathol.* 45, pp. 129–151. ISSN: 0066-4286.

Szekeres, Silvia, Mira Dauti, Caroline Wilde, Didier Mazel, and Dean A. Rowe-Magnus (2007). "Chromosomal toxin-antitoxin loci can diminish large-scale genome reduc-

tions in the absence of selection". In: *Molecular Microbiology* 63.6, pp. 1588–1605. ISSN: 1365-2958.

Tam, Jeffrey E. and Bruce C. Kline (1989). "The F plasmid ccd autorepressor is a complex of CcdA and CcdB proteins". In: *Molecular and General Genetics MGG* 219.1-2, pp. 26–32. ISSN: 0026-8925.

Tan, Qian, Naoki Awano, and Masayori Inouye (2011). "YeeV is an *Escherichia coli* toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB". In: *Molecular Microbiology* 79.1, pp. 109–118. ISSN: 1365-2958.

Tao, T., J. C. Bourne, and R. M. Blumenthal (1991). "A family of regulatory genes associated with type II restriction-modification systems". In: *Journal of Bacteriology* 173.4, pp. 1367–1375. ISSN: 0021-9193.

Taylor, Diane E. and Robert B. Grant (1976). "Inhibition of bacteriophage lambda, T1, and T7 development by R plasmids of the H incompatibility group". In: *Antimicrobial agents and chemotherapy* 10.4, p. 762.

Thomas, Christopher M. and Kaare M. Nielsen (2005). "Mechanisms of, and barriers to, horizontal gene transfer between bacteria". In: *Nature Reviews Microbiology* 3.9, pp. 711–721. ISSN: 1740-1526.

Toleman, Mark A. and Timothy R. Walsh (2011). "Combinatorial events of insertion sequences and ICE in Gram-negative bacteria". In: *FEMS microbiology reviews* 35.5, pp. 912–935. ISSN: 1574-6976.

Torres, Begoña, Susanne Jaenecke, Kenneth N Timmis, José L García, and Eduardo Díaz (2003). "A dual lethal system to enhance containment of recombinant microorganisms". In: *Microbiology* 149.12, pp. 3595–3601.

Tripathi, Arti, Pooja C Dewan, Bipasha Barua, and Raghavan Varadarajan (2012). "Additional role for the ccd operon of F-plasmid as a transmissible persistence factor". In: *Proceedings of the National Academy of Sciences* 109.31, pp. 12497–12502.

Tsilibaris, V., G. Maenhaut-Michel, N. Mine, and L. Van Melderen (2007). "What is the benefit to *Escherichia coli* of having multiple toxin-antitoxin systems in its genome?" In: *Journal of Bacteriology* 189.17, p. 6101.

Turner, Paul E. and Lin Chao (1998). "Sex and the evolution of intrahost competition in RNA virus φ6". In: *Genetics* 150.2, pp. 523–532. ISSN: 0016-6731.

Úbeda, Carles, Elisa Maiques, Erwin Knecht, Íñigo Lasa, Richard P Novick, and José R Penadés (2005). "Antibiotic–induced SOS response promotes horizontal

dissemination of pathogenicity island–encoded virulence factors in staphylococci". In: *Molecular Microbiology* 56.3, pp. 836–844. ISSN: 1365-2958.

Ul'yanova, V. V., V. I. Vershinina, M. A. Kharitonova, and M. R. Sharipova (2007). "The effect of Spo0A and AbrB proteins on expression of the genes of guanyl-specific ribonucleases from *Bacillus intermedius* and Bacillus pumilus in *Bacillus subtilis* recombinant strains". In: *Microbiology* 76.5, pp. 563–568. ISSN: 0026-2617.

Ulyanova, Vera, Valentina Vershinina, and Olga Ilinskaya (2011). "Barnase and binase: twins with distinct fates". In: *FEBS Journal* 278.19, pp. 3633–3643. ISSN: 1742-4658.

Unoson, Cecilia and E. Gerhart H. Wagner (2008). "A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli*". In: *Molecular Microbiology* 70.1, pp. 258–270. ISSN: 1365-2958.

V Znamenskaya, Lilia, Olga A Vershinina, Valentina I Vershinina, Inna B Leshchinskaya, and Robert W Hartley (1999). "Expression of the genes for guanyl-specific ribonucleases from *Bacillus intermedius* and *Bacillus pumilus* is regulated by the two component signal transduction system PhoP-PhoR in *B. subtilis*". In: *FEMS microbiology letters* 173.1, pp. 217–222.

Van Melderen, L. and M. S. De Bast (2009). "Bacterial Toxin-Antitoxin Systems: More Than Selfish Entities?" In: *PLoS Genetics* 5.3.

Van Melderen, Laurence (2010). "Toxin-antitoxin systems: why so many, what for?" In: *Current opinion in microbiology* 13.6, pp. 781–785. ISSN: 1369-5274.

Van Melderen, Laurence, Minh Hoa Dao Thi, Paolo Lecchi, Susan Gottesman, Martine Couturier, and Michael R. Maurizi (1996). "ATP-dependent Degradation of CcdA by Lon Protease EFFECTS OF SECONDARY STRUCTURE AND HETEROLOGOUS SUBUNIT INTERACTIONS". In: *Journal of Biological Chemistry* 271.44, pp. 27730–27738. ISSN: 0021-9258.

Vaughn, Jeffrey L., Victoria Feher, Stephen Naylor, Mark A. Strauch, and John Cavanagh (2000). "Novel DNA binding domain and genetic regulation model of *Bacillus subtilis* transition state regulator abrB". In: *Nature structural & molecular biology* 7.12, pp. 1139–1146.

Vazquez-Laslop, Nora, Hyunwoo Lee, and Alexander A. Neyfakh (2006). "Increased persistence in *Escherichia coli* caused by controlled expression of toxins or other unrelated proteins". In: *Journal of Bacteriology* 188.10, pp. 3494–3497. ISSN: 0021-9193.

Vega, Nicole M., Kyle R. Allison, Ahmad S. Khalil, and James J. Collins (2012). "Signaling-mediated bacterial persister formation". In: *Nature chemical biology* 8.5, pp. 431–433. ISSN: 1552-4450.

Velappan, Nileena, Daniele Sblattero, Leslie Chasteen, Peter Pavlik, and Andrew R. M. Bradbury (2007). "Plasmid incompatibility: more compatible than previously thought?" In: *Protein Engineering Design and Selection* 20.7, pp. 309–313. ISSN: 1741-0126.

Vogel, Jørg, Liron Argaman, E. Gerhart H. Wagner, and Shoshy Altuvia (2004). "The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide". In: *Current Biology* 14.24, pp. 2271–2276. ISSN: 0960-9822.

Wang, Ting, Sanja Tomic, Razif R. Gabdoulline, and Rebecca C. Wade (2004). "How optimal are the binding energetics of barnase and barstar?" In: *Biophysical Journal* 87.3, pp. 1618–1630. ISSN: 0006-3495.

Wang, Xiaoxue, Dana M. Lord, Hsin-Yao Cheng, Devon O. Osbourne, Seok Hoon Hong, Viviana Sanchez-Torres, Cecilia Quiroga, Kevin Zheng, Torsten Herrmann, and Wolfgang Peti (2012). "A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS". In: *Nature chemical biology* 8.10, pp. 855–861. ISSN: 1552-4450.

Wang, Xiaoxue, Dana M. Lord, Seok Hoon Hong, Wolfgang Peti, Michael J. Benedik, Rebecca Page, and Thomas K. Wood (2013). "Type II toxin/antitoxin MqsR/MqsA controls type V toxin/antitoxin GhoT/GhoS". In: *Environmental microbiology*. ISSN: 1462-2920.

Wang, Zhijun, Li Jin, Zhenghong Yuan, Grzegorz Węgrzyn, and Alicja Węfôgrzyn (2009). "Classification of plasmid vectors using replication origin, selection marker and promoter as criteria". In: *Plasmid* 61.1, pp. 47–51. ISSN: 0147-619X.

Washietl, Stefan, Sven Findeiß, Stephan A Müller, Stefan Kalkhof, Martin von Bergen, Ivo L Hofacker, Peter F Stadler, and Nick Goldman (2011). "RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data". In: *RNA* 17.4, pp. 578–594. ISSN: 1355-8382.

Weaver, Keith E. (2012). "The par toxin-antitoxin system from Enterococcus faecalis plasmid pAD1 and its chromosomal homologs". In: *RNA biology* 9.12, pp. 15–14. ISSN: 1547-6286.

Weaver, Keith E., Kristi D. Jensen, Amy Colwell, and Saileela Sriram (1996). "Functional analysis of the Enterococcus faecalis plasmid pAD1-encoded stability determinant par". In: *Molecular Microbiology* 20.1, pp. 53–63. ISSN: 1365-2958.

Weaver, Keith E., Shirisha G. Reddy, Cassandra L. Brinkman, Smita Patel, Kenneth W. Bayles, and Jennifer L. Endres (2009). "Identification and characterization of a family of toxin-antitoxin systems related to the Enterococcus faecalis plasmid pAD1 par addiction module". In: *Microbiology* 155.9, pp. 2930–2940. ISSN: 1350-0872.

Weaver, Keith E., Dariel M. Weaver, Carol L. Wells, Christopher M. Waters, Marshall E. Gardner, and Erik A. Ehli (2003). "Enterococcus faecalis plasmid pAD1-encoded Fst toxin affects membrane permeability and alters cellular responses to lantibiotics". In: *Journal of Bacteriology* 185.7, pp. 2169–2177. ISSN: 0021-9193.

Wei, Hao, Richard Meilan, Amy M. Brunner, Jeffrey S. Skinner, Caiping Ma, Harish T. Gandhi, and Steven H. Strauss (2007). "Field trial detects incomplete barstar attenuation of vegetative cytotoxicity in Populus trees containing a poplar LEAFY promoter:: barnase sterility transgene". In: *Molecular breeding* 19.1, pp. 69–85. ISSN: 1380-3743.

Werner, Finn and Dina Grohmann (2011). "Evolution of multisubunit RNA polymerases in the three domains of life". In: *Nature Reviews Microbiology* 9.2, pp. 85–98. URL: http://www.nature.com/nrmicro/journal/v9/n2/abs/nrmicro2507.html;%20http://www.researchgate.net/publication/49754339_Evolution_of_multisubunit_RNA_polymerases_in_the_three_domains_of_life/file/9fcfd5087ed0867962.pdf.

WHO (2014). *Antimicrobial Resistance: Global Report on surveillance.* World Health Organization.

Wilbaux, M., N. Mine, A. M. Guerout, D. Mazel, and L. Van Melderen (2007). "Functional interactions between coexisting toxin-antitoxin systems of the ccd family in *Escherichia coli* O157: H7". In: *Journal of Bacteriology* 189.7, p. 2712.

Williams, Julia J. and Paul J. Hergenrother (2012). "Artificial activation of toxin–antitoxin systems as an antibacterial strategy". In: *Trends in microbiology* 20.6, pp. 291–298. ISSN: 0966-842X.

Willms, Allan R., Paul D. Roughan, and Jack A. Heinemann (2006). "Static recipient cells as reservoirs of antibiotic resistance during antibiotic therapy". In: *Theoretical population biology* 70.4, pp. 436–451. ISSN: 0040-5809.

Wit, Pierre, Melissa H. Pespeni, Jason T. Ladner, Daniel J. Barshis, François Seneca, Hannah Jaris, Nina Overgaard Therkildsen, Megan Morikawa, and Stephen R. Palumbi (2012). "The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis". In: *Molecular ecology resources* 12.6, pp. 1058–1067. ISSN: 1755-0998.

Wozniak, Rachel A. F. and Matthew K. Waldor (2009). "A toxin-antitoxin system promotes the maintenance of an integrative conjugative element". In: *PLoS Genetics* 5.3, e1000439. ISSN: 1553-7404.

Yahara, Koji, Mikihiko Kawai, Yoshikazu Furuta, Noriko Takahashi, Naofumi Handa, Takeshi Tsuru, Kenshiro Oshima, Masaru Yoshida, Takeshi Azuma, and Masahira Hattori (2012). "Genome-wide survey of mutual homologous recombination in a highly sexual bacterial species". In: *Genome biology and evolution* 4.5, pp. 628–640. ISSN: 1759-6653.

Yakovlev, G. I., G. P. Moiseyev, I. I. Protasevich, B. Ranjbar, A. L. Bocharov, M. P. Kirpichnikov, R. M. Gilli, C. M. Briand, R. W. Hartley, and A. A. Makarov (1995). "Dissociation constants and thermal stability of complexes of *Bacillus intermedius* RNase and the protein inhibitor of *Bacillus amyloliquefaciens* RNase". In: *Febs Letters* 366.2-3, pp. 156–158.

Yamaichi, Yoshiharu, Michael A. Fogel, and Matthew K. Waldor (2007). "par genes and the pathology of chromosome loss in *Vibrio cholerae*". In: *Proceedings of the National Academy of Sciences* 104.2, pp. 630–635. ISSN: 0027-8424.

Zhang, Junjie, Yonglong Zhang, and Masayori Inouye (2003). "Characterization of the interactions within the *mazE*F addiction module of Escherichia coli". In: *Journal of Biological Chemistry* 278.34, pp. 32300–32306. ISSN: 0021-9258.

Zhang, Yi Xuan, Xiao Kui Guo, W. U. Chuan, B. I. Bo, Shuang Xi Ren, Chun Fu Wu, and Guo Ping Zhao (2004). "Characterization of a novel toxin-antitoxin module, VapBC, encoded by *Leptospira interrogans* chromosome". In: *Cell research* 14.3, pp. 208–216. ISSN: 1001-0602.

Zhang, Yonglong, Junjie Zhang, Klaus P. Hoeflich, Mitsuhiko Ikura, Guoliang Qing, and Masayori Inouye (2003). "MazF Cleaves Cellular mRNAs Specifically at ACA to Block Protein Synthesis in *Escherichia coli*". In: *Molecular cell* 12.4, pp. 913–923. ISSN: 1097-2765.

Znamenskaya, L. V., L. A. Gabdrakhmanova, E. B. Chernokalskaya, I. B. Leshchinskaya, and R. W. Hartley (1995). "Phosphate regulation of biosynthesis of extra-

cellular RNases of endospore-forming bacteria". In: *Febs Letters* 357.1, pp. 16–18.

Znamenskaya, L. V., M. A. Kharitonova, A. R. Kayumov, and S. I. Krasnov (2002). "The biosynthesis of new secretory high-molecular-weight ribonucleases in *Bacillus intermedius* and *Bacillus subtilis*". In: *Microbiology* 71.6, pp. 690–697. ISSN: 0026-2617.